AFIT/GST/ENS/93M-05

AD-A266 719

DTIC
S ELECTE
JUL 8 1993
C D

AN ANALYSIS OF TACTICAL AIRLIFTER
CHARACTERISTICS AND THEIR IMPACT ON
THEATER AIRLIFT SYSTEM PERFORMANCE

THESIS

John J. Koger, Captain, USAF

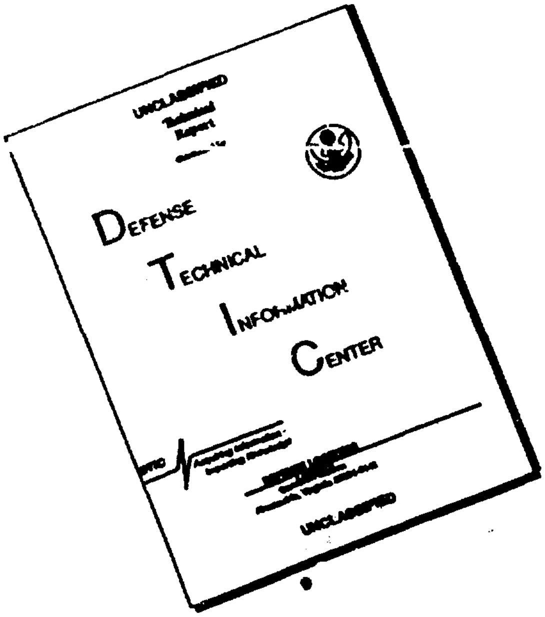
AFIT/GST/ENS/93M-05

Approved for public release; distribution unlimited

93-15349

124pg

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

AFIT/GST/ENS/93M-05

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

DTIC QUALITY INSPECTED 5

Accesion For		
NTIS	CRA&I	☑
DTIC	TAB	☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

AFIT/GST/ENS/93M-05

AN ANALYSIS OF TACTICAL AIRLIFTER CHARACTERISTICS AND THEIR IMPACT ON THEATER AIRLIFT SYSTEM PERFORMANCE

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Strategic and Tactical Sciences

John J. Koger, B.S.

Captain, USAF

March 1993

Approved for Public release; distribution unlimited

Preface

This research addresses the issue of which tactical airlifter characteristics most influence theater airlift system performance. By evaluating specifically selected combinations of airlifter characteristics, their effect on theater airlift system performance is assessed. Measures of system-level performance are provided by the Generalized Air Mobility Model (GAMM), an established simulation of a theater airlift system.

I decided on this topic after reading an Air Force Institute of Technology thesis by Royal Australian Air Force Flight Lieutenant Paul Pappas. Using two analysis techniques of great interest to me, simulation and response surface methodology, he analyzed the impact of varying airlifter characteristics in a Central American scenario. This research augments his work by considering a more complete list of characteristics in the vastly different operating environment of Southwest Asia. In combination, the two studies should provide insight into which characteristics are significant in general, and which characteristics are scenario dependent.

I owe a debt of gratitude to my advisor, Lt Col Paul Auclair, for his continued support on this project. His experience and knowledge in simulation, response surface methodology, and statistics proved invaluable. On more than one occasion, he brought up important considerations that might have otherwise gone unnoticed. Steven J. Wourms, Chief of Assessments Section, Studies and Analysis Division of the Development Planning Directorate of the Aeronautical Systems Center at Wright-Patterson AFB, was also very helpful. He was always available to answer my questions on GAMM, and his tactical airlift analysis experience was insightful. I would also like to thank my reader, Maj Dennis Dietz, a respected operations analyst with tactical airlift experience.

Finally, I would like to thank my wife, Tara, and my children, Ashley and Ryan, for putting up with me for the six months I worked on this project.

John Koger

Table of Contents

List of Figures

List of Tables

Table

AFIT/GST/ENS/93M-05

Abstract

This study used computer simulation and Response Surface Methodology to determine what tactical airlifter characteristics most impact theater airlift performance in a Southwest Asia (SWA) scenario. Some aircraft characteristics were grouped into functional sets, while others were considered individually. After screening one characteristic, reliability, with a two-level factorial experiment, a Box-Behnken design was used to estimate second-order metamodels. A stepwise regression procedure indicated that, if attrition rates are ignored, airlift system performance is most impacted by the aircraft's size, survivability, cruise speed, and ability to operate on short fields. The SWA scenario used in this study covers a large geographical area and varying threat levels and types. The results of this study were compared with those of an earlier study that used the much smaller, low threat Central American scenario. It was determined that across a range of scenarios airlift system performance is most affected by the aircraft's size, survivability, cruise speed, ability to operate on short fields, and ability to operate on unprepared surfaces.

AN ANALYSIS OF TACTICAL AIRLIFTER CHARACTERISTICS
AND THEIR IMPACT ON THEATER AIRLIFT SYSTEM PERFORMANCE

I. Introduction

Background

Considerable research has been done by the Air Force, the Army, and major aircraft manufacturers concerning future theater airlift requirements. The most comprehensive work is the ongoing combined effort of Air Mobility Command (AMC) and the Air Force Aeronautical Systems Center (ASC) under the heading of Future Theater Airlift Studies (FTAS). This chapter will provide a brief background on theater airlift, FTAS, and the 1991 AFIT thesis this research augments.

Tactical Airlift. Strategic airlift involves moving personnel, supplies, and equipment, to, from, or between theaters of operations. Theater airlift involves such movements within a theater of operations. The USAF defines five general mission categories for tactical airlift (theater and tactical can be used interchangeably in this context):

1. Deployment is the movement of forces to their initial area of operation in a theater.
2. Employment is the movement of forces within a theater after their initial deployment.
3. Sustainment is the movement of replacement personnel and supplies in support of deployed forces.
4. Retrograde is the use of aircraft on the return leg of a deployment mission, such as for evacuees or enemy prisoners of war.

5. Air Evacuation is the air transport of casualties to suitable medical facilities. This mission usually requires a specially configured aircraft (20:8-9).

The fundamental objective of theater airlift is to provide the user, which is usually the Army, with the capability to conduct a cargo movement mission when there are no other means available. Theater airlift allows the Army increased mobility without reliance upon "ground lines of communication." It affords the ability to maneuver when the lack of roads, rail lines, or major airfields might otherwise preclude it (20:5-6).

The Airland Battle. The US Army continues to develop Airland Operations (ALO), its newest war fighting doctrine. Under ALO, the concept of the battlefield has drastically changed (17:1139). Instead of a single line on the map designating the Forward Line of Troops (FLOT), the airland battle will consist of non-linear battlefields. Lighter, more mobile units will carry out deep attacks on second echelon units while others simultaneously engage first echelon units. Those US forces carrying out the deep attacks will be heavily dependent upon the theater airlift system for their initial employment, resupply, and extraction. Such high threat operations into austere locations indicate a need for a survivable airlifter capable of loading and unloading very quickly, without external support.

General Issue. The primary tactical airlifter in Air Mobility Command (AMC) is the C-130 Hercules, which became operational in the 1950s. Although it is still an effective tactical airlifter, the C-130 is based on a 40 year old design that did not anticipate the significantly changed environment that tactical airlifters are likely to face. The increased threat dictated by the Army's ALO highlights the vulnerability of the C-130, which has a wet wing and no provisions for

radar warning receivers, electronic countermeasure pods, or dispensers for flares or chaff (10:43-44). A tactical airlifter suited to the demands of the battlefield environment should be considered to replace the C-130. It should include survivability and vulnerability improvements to operate in a modern, high-threat environment, and use new technology to improve upon the C-130's performance and cargo handling capabilities.

Future Airlift Theater Studies. Many of the C-130s in use today need to be replaced or refurbished as they are approaching the end of their useful service life. The C-17, a large transport suitable for theater airlift, will soon be added to the inventory. Although the C-17 will be capable of short field operations, it is intended primarily to augment the C-5 and C-141 strategic airlift fleet, and occasionally deliver critical cargo directly into forward operating locations. To assist Headquarters AMC in assessing theater airlift systems of the future, the Aeronautical Systems Center (ASC) is conducting the most comprehensive analysis of theater airlift capability ever carried out by the US Air Force through the Future Theater Airlift Studies (FTAS) project. The objective of FTAS is to establish the analytical basis to support Headquarters AMC in the development of a Mission Needs Statement (MNS) for the next generation USAF tactical airlifter. The MNS is required to initiate a Department of Defense (DOD) major weapon system acquisition process.

Thesis Review. In 1991, Paul Pappas, a Flight Lieutenant in the Royal Australian Air Force, completed a thesis for the Master of Science in Logistics Management program at AFIT. His research, "An Analysis of Fixed Wing Tactical Airlifter Characteristics Using an Intra-Theater Airlifter Computer Model", used a computer simulation model developed for FTAS and *response surface methodology* to determine which tactical

airlifter characteristics most significantly affected tactical airlift capability in a particular scenario. The theater considered covered a small geographical area with many short, unimproved airfields. The baseline tactical airlifter was the C-130H, and airlifter characteristics were aggregated into common groups to provide ease of analysis. The airlifter characteristics were systematically varied about the C-130H baseline to determine which of them significantly affect the capability of the tactical airlift system in this scenario. Pappas's research indicated that airlift system performance, for the scenario considered, was influenced by deviations from the baseline C-130H for only two groups of characteristics: the size of the aircraft's cargo bay and the aircraft's ability to operate on unprepared surfaces (16:91).

Prior to the Pappas research ASC had been unable to successfully apply experimental design to study large numbers of tactical airlifter characteristics. Pappas's work established that, when combined into functional groups, large numbers of characteristics can be successfully analyzed using experimental design.

Knowing the results of Pappas's experiment facilitates much more research in this area. Specifically, he analyzed the performance of the tactical airlifter in a single scenario. Since it seems improbable that a new airlifter will be designed for just this case, there is a need to analyze tactical airlifter characteristics across a range of operating conditions, or scenarios. With additional, much different, scenarios assessed, it would be possible to identify tactical airlifter characteristics that result in an increased throughput across a range of operating environments.

Additionally, knowing Pappas's results allows similar analysis to be conducted on those characteristics he did not study. In the Pappas experiment, such presumably important characteristics as

maintainability, reliability, and survivability were not considered. An analysis that included these characteristics, as well as those studied by Pappas, would more completely determine which airlifter characteristics most impact tactical airlift capability.

Research Objective

Using the techniques of response surface methodology and computer simulation this research will investigate how aircraft characteristics affect tactical airlift capability. Given a specific theater of operations that covers a relatively large geographical area and a specified set of airlift requirements, answers to the following questions will be sought:

1. Which aircraft characteristics adequately characterize tactical airlift system performance?
2. Are traditional strategic airlift measures appropriate when measuring the capability of a tactical airlift system that covers a large geographical area?
3. Which aircraft characteristics prove significant across a range of operating conditions and which characteristics are only significant in certain scenarios?

Scope

This research is limited to a study of tactical airlifter characteristics using the simulation model and scenario described in Chapter II. Both the model and the scenario have been accepted as very useful analytical tools within the theater airlift analysis community. No verification or validation of the scenario or model will be attempted in this thesis. However, previous verification and validation efforts will be discussed.

Summary

This chapter defined tactical airlift and the role it will play in future battles according to the Army's Airland Operations (ALO). It described how the 40 year old C-130, the USAF's primary tactical airlifter, was never designed to meet the increased demands of the Army's ALO concept. It then described how the USAF is exploring the need for a follow on tactical airlifter through Future Theater Airlift Studies (FTAS).

Finally, an earlier analysis of aircraft characteristics and their impact on airlift system performance was reviewed. This brief review pointed to the need for more research in this area. The research objectives presented are intended to address many of the open questions concerning the effect of airlifter characteristics on tactical airlift capability.

II. Generalized Air Mobility Model (GAMM)

Introduction

GAMM is a detailed simulation model that serves as the primary analytical tool for Future Theater Airlift Studies (FTAS). The two main theater airlift simulation models in existence prior to FTAS, the Tactical Airlift System Comparative Analysis Model (TASCAM) and the Scenario Unrestricted Mobility Model for Intra-Theater Simulation (SUMMITS), were based on a *wholesale* airlift system. In a wholesale system, cargo is delivered to centralized locations for further dispersal. Recent emphasis, as outlined in the Army's Airland Operations (ALO), is on a much more *retail* theater airlift system. In a retail system, cargo is delivered directly to the end user. Future airlifters, operating within a retail theater airlift system, will face a denser, more capable threat environment and more austere operating conditions with very limited cargo handling support. Such operations will also entail closer work with the Army and forward based Air Force fighter units (6:2-3).

As the salient features of retail theater airlift operations are not accounted for by TASCAM and SUMMITS, the Aeronautical System Center (ASC) sponsored development of the Generalized Air Mobility Model (GAMM) by the General Research Corporation (GRC). GAMM is a detailed Monte-Carlo simulation of a theater airlift system. The model was developed for the Digital Equipment Corporation (DEC) VAX line of computers and is written in the Simscript II.5 simulation language and the Fortran 77 programming language. As a minimum configuration GAMM requires a MicroVax II with 6 Megabytes of random access memory and 50 Megabytes of disk storage capacity (8:5).

The Primary Simulation Elements

Figure 1 illustrates the main elements of the GAMM model. These elements include the GAMM inputs, the GAMM program, and the GAMM outputs. The GAMM inputs consist of the Scenario File, the Jobs File and the Runtime Command File.

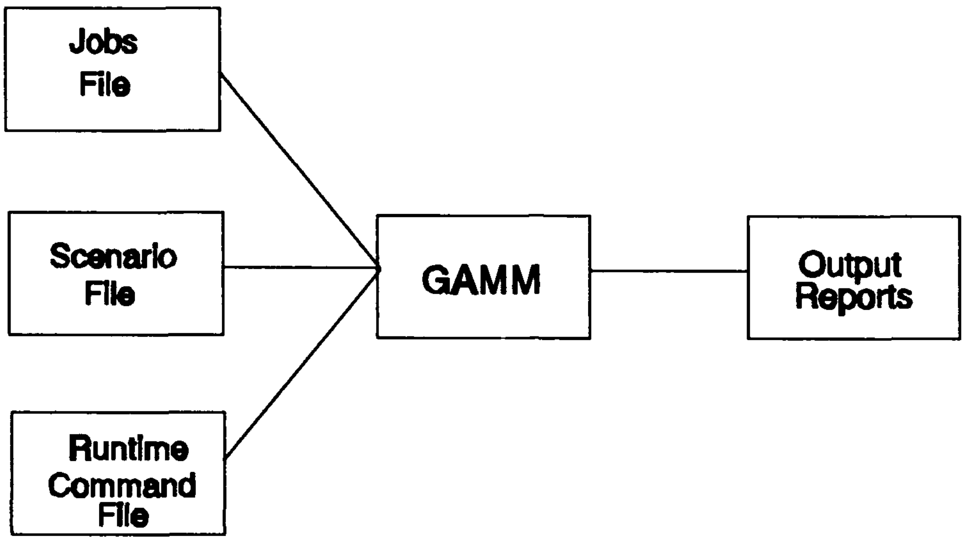


Figure 1. Elements of GAMM

The Scenario File describes the environment being modeled, including airbases and entry/delivery sites. Airbase information provides the location, identification, and descriptive data relative to each airbase. Entry/delivery site information describes the name and location of each site where airlift jobs originate and/or terminate. In addition, each entry/delivery site has information identifying airbases which serve that entry/delivery site. The Scenario File also describes each airlifter type's capability to transport cargo in terms of runway length requirements and weight limitations, speed, cargo compartment size, and reliability/maintainability characteristics. The Jobs File describes the cargo that will be flowing through the transportation system. It includes a description of the job request, an entry site, a delivery site, an entry time, a delivery time and the job priority. Dimensional data and weight are also listed. Finally, the Runtime Command File contains operational and analysis conditions for the simulation. (8:4)

The GAMM Output Reports include many different hardcopy analysis products such as airlifter flight histories, scenario data, cargo movement statistics, and various measures of effectiveness. The *General Summary Report* is the primary output product and is a comprehensive summary of a specific GAMM run.

GAMM as a Transportation System

In GAMM, the required cargo movements for a scenario are established prior to starting the simulation. In other words, cargo movement requirements are known prior to starting the war, and do not change as the war progresses. Modeling the tactical airlift system in this manner removes the complicating effects of combat modeling and simplifies the primary objective of measuring airlift system performance (6:3).

A transportation system can be thought of as a network in which an item that enters at an entry site is moved, via inter-connected arcs and nodes, to a delivery site. Figure 2 illustrates the network structure used in GAMM. The entry/delivery (E/D) sites and airbases are represented as nodes. These nodes are connected by two types of arcs, which are distinguished by mode of transportation. The air transportation arcs use airlifter assets, while the transhipment arcs do not (8:3-1). Henceforth the term "transhipment" will be used to refer to cargo movement from the entry site to the originating airbase and from the receiving airbase to the delivery site. Thus, transhipment refers to cargo movement, via non-airlifter assets, along the first and third legs of the transportation system depicted in Figure 2 (8:3-2).

GAMM Overview

When the GAMM simulation begins, all airlifters are at their home bases. When an airlift job enters the transportation system at an entry site, it is transported via a transhipment arc to an airbase associated

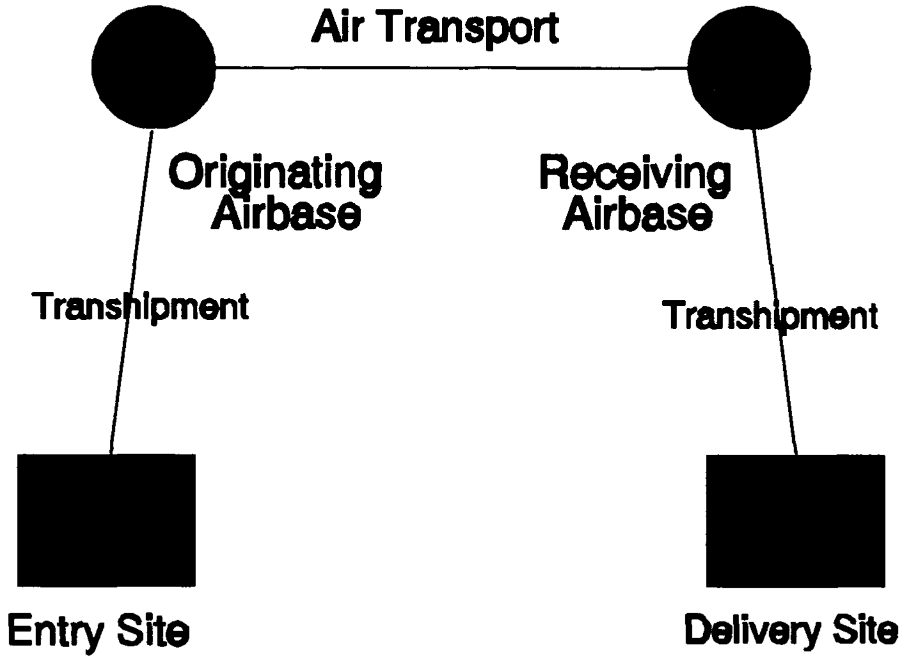


Figure 2. GAMM as a Transportation System (8:Sec 3-1)

with that entry site. For originating airbases with available airlifters, jobs are scheduled to depart based on the priority and the weight of the cargo. There are nine cargo priorities, 1 to 9, with priority 1 being the highest. The significance of a priority number is established by the *priority scheduling factor*, which is set in the Runtime Command File. Jobs with a priority number higher than the priority scheduling factor will be scheduled in order of their priority with ties being scheduled according to weight. All other jobs are considered equal in priority. When this is the case, jobs are scheduled according to weight, from heaviest to lightest (8:3-2).

Prior to loading the cargo on an airlifter, GAMM checks each receiving airbase associated with the delivery site for sufficient runway length, runway load bearing limitations, fuel availability, and

maximum aircraft on the ground limitations (MOG). The airfields are checked in the order of preference specified in the Scenario File for the transhipment links to each E/D site. If there is not a suitable airfield for delivery, there are three possibilities. First, a search for *opportune sites* that are linked to the delivery site is made. Opportune sites are defined as austere or unimproved airland sites which can support airlift operations for a limited time (6:46). Examples of opportune sites include roads, cleared fields, or helicopter pads. Next, the model will schedule an airdrop if an airdrop delivery is acceptable for the cargo. The third possibility is to block the cargo from further progress because no feasible link to the delivery site exists. Finally, useful load is determined prior to an aircraft being relocated to or loaded at an originating airbase. If no load can be carried, that cargo is blocked because it cannot be moved by the airlifter (8:3-2 to 3-3); GAMM does not allow cargo to be transhipped directly from its entry site to its delivery site.

If GAMM determines that a load can be delivered, loading commences on the scheduled aircraft. Two loading methods are provided: weight-and-volume loading or weight-only loading. The user determines the loading method at run time. For scenarios with a wide variety of cargo and vehicles, the weight-and-volume method is more realistic and is recommended (8:3-3).

Combat and Attrition. GAMM models the combat aspects of the scenario through the specification of probabilities of survival inflight and on the ground. Probabilities of battle damage inflight and on the ground are also modeled. The survival and battle damage probabilities for aircraft on the ground are currently set to unity pending the outcome of efforts that will yield realistic estimates for these values (6:38). In essence, GAMM currently ignores damage and attrition to aircraft on the ground. Inflight probabilities are based on the

airlifter's distance from the Forward Line of Troops (FLOT). Table 1 displays the values contained in the Scenario File for the probability of survival and the conditional probability of no battle damage, given aircraft survival. These probabilities are airbase attributes that can take on different values for different airlifter types. The values listed in Table 1 are currently applied to the C-130H. To summarize the information in Table 1, the closer the aircraft is to the FLOT the more likely it will be killed or damaged.

Table 1. GAMM Probability of Survival Factors (6:64)

Minimum Distance to FLOT (km)	Maximum Distance to FLOT (km)	Probability of Survival Factor	Conditional Probability No Battle Damage
10	9999	1.000	1.000
0	10	.983	.995
-50	0	.969	.915
-9999	-50	.970	.918

Inflight Survivability. Inflight survivability is a function of the airlifter type, the originating airbase, and the receiving airbase (6:63). The equation for inflight survivability is as follows:

$$P_s(F) = P_s(O)*P_s(R) \qquad (1)$$

where

$P_s(F)$ is the probability of survival on a flight between the originating and receiving airbases of interest for the given airlifter type,

$P_s(O)$ is the probability of survival factor at the originating airbase for the given airlifter type, and

$P_s(R)$ is the probability of survival factor at the receiving airbase for the given airlifter type.

When the aircraft lands, a random draw is made. If that random draw is greater than $P_s(F)$, the airlifter was destroyed inflight.

If an aircraft was carrying cargo when it was destroyed, the cargo that is destroyed is reordered. The resulting delay may cause a late delivery or, in some cases, deletion of the job request. Aircraft that are destroyed during the simulation are not replaced.

This has been a brief overview of how a theater airlift system is modeled by GAMM. For more detailed information the interested reader should refer to the *GAMM Programmer Analyst Manual* (8).

GAMM Scenarios

Three scenarios presently exist for GAMM; Europe (EUR), Southwest Asia (SWA), and Central America (CA). While their originating context is outdated, the scenarios continue to be used within the theater airlift analysis community because they cover a good mix of theater characteristics such as: combat intensity, threat level, flight distances, pressure/altitude requirements, airfields, cargo, and on-ground infrastructure (17:1140). For example, the EUR scenario is characterized by a high intensity combat environment, a cool and wet climate, a well established transportation infrastructure, and rolling plain terrain features. In contrast, the CA scenario is characterized by light-to-mid-intensity combat, a wet and warm to hot climate, a primitive transportation infrastructure, and mountainous terrain (7:D-2 - D-3).

Scenario Selection

A previous thesis effort performed an analysis of airlifter characteristics similar to that conducted in this research. That analysis used GAMM's Central America (CA) scenario, which covers a small geographical region with few roads and almost no railway system. Although this scenario includes a limited number of major airfields,

there are many short, unprepared, grass or dirt airfields. The Southwest Asia scenario was selected for this research. It covers a much larger area, a more established infrastructure, and a considerable number of major airfields and short, unprepared landing surfaces. By choosing a scenario that is vastly different than that used previously, the results of the two studies can be compared to make some generalizations about significant airlifter characteristics.

Scenario Assumptions

The Southwest Asia Scenario is based on the following assumptions:

1. The scenario occurs between 2000 and 2010.
2. It is a conventional war where both sides abide by the provisions of the INF Treaty.
3. Army Airland Operations are used. They represent the fighting doctrine now being developed by the Army for a conventional war.
4. Air Force 2000 concepts are used.
5. A 30 day war is postulated to ensure logistics resupply tasks are examined, since prepositioned stocks may suffice for a short conflict and not fully stress an airlift force. Also, the conflict must be long enough to reveal the full range of intratheater requirements (6:5-6).

Southwest Asia Scenario Overview

The conflict in this scenario arises from a Soviet drive, through Iran, toward the oil fields of the Middle East. The war is characterized by combat intensity that is high at times, although not sustained for the entire 30 days. Although there is a large number of enemy forces, due to the vast area of operations, the forward line of troops (FLOT) is discontinuous. The climate is essentially dry and temperatures vary greatly between the two primary terrain features:

mountain ranges that peak at 8,000-19,000 feet and the desert like Central Plateau. The area of operations is approximately 850 by 900 miles, or about three times the size of France. The war begins to wind down by the twenty-first day of fighting as rising tensions in other parts of the world relegate the Iranian front to a minor theater for both main combatants. After the twenty-first day of fighting, both sides are primarily involved with pulling back into defensive positions (7:D-2).

Southwest Asia Airlift Job Definitions

The U.S. agrees to assist Iran and begins deploying intratheater airlifters to the Persian Gulf region within three days of the invasion. The land component deployed consists of two airborne and one infantry division, an air cavalry brigade, and a Marine Expeditionary Brigade. In addition, eight fighter wings are moved to the region. For the SWA scenario, 31 representative jobs were identified. The details of those jobs are presented in Appendix A. Almost 43,000 tons of cargo require movement over the 30 day period.

In closing this section, it should be noted that considerable effort went into selecting the specific airlift jobs used in this scenario.

> The job selection process was guided by the scenario, historical analysis, extensive research, and the professional experience of analysts. Doctrinal statements provided the framework for postulated tactical and operational level options which generated airlift requirements. A working group comprised of Air Force, Army, and airframe contractor representatives has greatly assisted in these airlift job definitions. (7:4-2)

GAMM Validation and Verification

Prior to using GAMM for this experiment, it was necessary to determine the extent to which the model was verified and validated. Validation can be defined as substantiation that a computerized model within its domain of applicability possesses a satisfactory range of

accuracy consistent with the intended application of the model (19:33). Model verification is usually defined as ensuring that the computer program of the computerized model and its implementation are correct (19:33). More simply, "validation deals with building the right model, verification deals with building the model right" (21:559). This section will outline the verification and validation that has been completed on GAMM Version 3.5 and the Southwest Asia Scenario database.

Verification of GAMM. Verification of a simulation model boils down to confirming that the computer code does what it was intended to do. Such confirmation is an essential step in the simulation model development process and should be thoroughly documented. The GAMM Programmer/ Analyst's Manual and the GAMM User's Manual both contain some very useful information relating to model verification (6:4-1 to 4-52; 8:2-16 to 2-31). Section 4 of the GAMM Programmer/Analyst Manual details the most significant events and routines within the model. Logic is described and flow charts are provided to illustrate actual coding. In Section 2 of the GAMM User's Manual a simple example of a tactical airlift system is used to step through the GAMM flight scheduling algorithm. The example first lists the rules that make up the basic scheduling algorithm and then provides a series of simple diagrams to illustrate the movement of aircraft and cargo within the system.

The most recent formal verification of GAMM was conducted in early 1992. Under TASK 0008 of the Future Theater Airlift Studies Southwest Asia Scenario: Generalized Air Mobility Model Data File Generation, the General Research Corporation and Ball Systems Engineering Division were contracted to conduct an extensive review of the data base used for the SWA scenario (6: Foreword). As a part of this review, verification of certain algorithms within the GAMM code was carried out. In particular,

verification of airlifter scheduler algorithms uncovered four deficiencies related to the allocation of airlifter resources.

> The correction of these deficiencies has reduced the total flight hours, sorties, and number of non-productive airlifter relocations while increasing the efficiency of the airlifter scheduling within the SWA scenario. (6:80)

Other improvements that resulted from the review conducted under Task 0008 included:

1. Development of a methodology to generate opportune sites at each delivery site.
2. The preferred run-time command file settings were changed.
3. Job deletion times were changed for late cargo.
4. Minor changes were made to runway attack and repair algorithms (6:86).

Validation of GAMM. As mentioned above, validation is concerned with building the right model. In general, there are two types of model validation; *conceptual* and *operational*. Conceptual validation is defined as determining that the theories and assumptions underlying the conceptual model are correct and that the model representation of the system is "reasonable" for the intended use of the model (19:33). Operational validation is defined as determining that the model's output behavior has sufficient accuracy for its intended purpose or use over the domain of the model's intended application (19:33).

Conceptual Model Validation. GAMM has evolved from a simplistic tactical airlift model to a more elaborate and more realistic model of a future tactical airlift system (8:1). It has gone from the original Version 0.5 to the current Version 3.5. Each version of GAMM represents a major change to the previous model. These major changes are a result of the ongoing conceptual model validation process. At each iteration, the conceptual model of the tactical airlift system was improved. For example, in going from Version 3.4 to 3.5, the major

change consisted of adding a logistical pipeline delay. This delay represented the amount of time it would take a Maintenance Repair Team to reach an aircraft that had broken down away from its home station. This change is a refinement to the conceptual model and makes it a more accurate representation of an actual tactical airlift system. The conceptual model validation of GAMM will continue until the level of accuracy considered necessary is obtained.

Operational Validation. When possible, operational validation is carried out by comparing model output data to system output data. If the system is not observable, the next choice is to compare model output with the output of other models of the system. In this case, the system is not observable, and no other comparable models exist. Thus, the operational validation process was limited to exploring model behavior. As a part of Task 0008, mentioned above, an extensive sensitivity analysis was conducted. The parameters that were varied included airlifter fleet size, airlifter home basing strategies, the average daily temperature, and the version of the SWA scenario being used (i.e. old, interim, or new). The findings of this analysis were summarized as follows.

> GAMM 3.5 has been successfully evaluated and tested during the execution of this task. Numerous anomalies were identified and corrected in the coding. These anomalies affected the treatment of cargo, the number of relocations, and other performance related metrics. The model has been fully screened and is ready to be used in an analysis of FTAS conceptual aircraft (6:125).

Data Validation. The primary purpose of Task 0008 was to validate the SWA scenario database. Together, the Jobs File and the Scenario File make up a particular GAMM scenario's database. In addition to validating these files, Task 008 reviewed each of the operational parameters that are used in the Runtime Command File.

Jobs File Validation. Before covering Jobs File validation, the actual Jobs File development process should be noted. The data used in the Jobs File was compiled and reviewed by a Jobs Working Group (JWG).

> The JWG is comprised of members from the Air Force acquisition, Air Force airlift, Army transportation, Army doctrine, Army aviation, and airframe manufacturer communities. The establishment of the JWG provided a forum for the interchange of information among the services and industry concerning the definition of a representative jobs set. This group, in conjunction with other working groups (such as the Cargo Handling Working Group), assisted in the definition of the assets, deployments, capabilities, and concept of operations to be depicted within the SWA scenario (7:1).

As a result of Task 0008, two major areas of the Jobs File were improved. First, a review of *job deletion times* was conducted. The job deletion time represents how late a job may be before it is deleted (7:32). This review resulted in the implementation of a new set of job deletion times that were more logically related to the particular job's priority (7:32).

The second improvement was related to method of delivery. Previously, GAMM would first attempt to use the standard airland method of delivery, even if an item was designated to be airdropped. Airdrop was only used by the simulation if airland was not feasible. This method was not acceptable for job types that could only be delivered by the air. As a result, changes were made to both the Jobs File and the Scenario File to ensure that cargo is delivered as initially intended (7:35).

Scenario File Validation. Much of the Task 0008 effort was dedicated to the validation of the SWA Scenario File. As a result major changes and enhancements were made to this file. Two of the major improvements are detailed below.

A review and comparison of data sources revealed airbase information included numerous errors (6:44). Errors in runway length,

airfield elevation, and runway material composition were not uncommon. The original source of airbase information was an unclassified, limited version the Defense Mapping Agency's (DMA) Automated Air Facilities Information File (AAFIF). To validate the DMA AAFIF data, General Research Corporation (GRC) accumulated data for comparison from Operational Navigational Charts, Tactical Pilotage Charts, the DOD Enroute Supplement, and the Air Mobility Command's Airfield Suitability Report. The source of any significant differences in the data was identified and changes to the original database were made when appropriate (6:46).

Another enhancement to GAMM that resulted from the Task 0008 effort was the addition of *opportune sites.* It was believed that "the incorporation of opportune sites would provide realistic alternative airland sites to permanent airbases, from both a ground and air commander's operational perspective" (6:46). Since examples of opportune sites include roads and cleared fields it can be inferred that the addition of opportune sites will add numerous transhipment links. They also allow the cargo to be airlifted to a site closer to the end customer.

Runtime Command File Validation. As a part of Task 0008, an extensive review of the operational parameters that make up the Runtime Command File was conducted. These parameter values are entered by the user at runtime and include the length of the simulation, the loading methods to be employed, the length of the crew duty day, and a number of other operational factors. This review provided a recommended value for each of the Runtime Command File parameters along with the rationale for why this value was selected.

The results of the Runtime Command File review are presented in Appendix B. It should be noted that the recommended values were used in the Runtime Command Files of this experiment.

Summary

This chapter provided a brief background of the Generalized Air Mobility Model (GAMM). First, the primary components of GAMM were discussed. It was shown that the inputs to GAMM include the Jobs File, the Scenario File, and the Runtime Command File. The other primary components included the GAMM program and numerous output reports.

Next, the way GAMM models a transportation system was covered. Within GAMM, cargo is transshipped from an entry site to an originating airbase via a transhipment arc. The cargo is next moved to a receiving airbase via an air transport arc. Finally, it is again transhipped from the receiving airbase to the delivery site.

An overview of how GAMM schedules cargo movements, or jobs, was also provided. Jobs are scheduled based on priority and weight. Prior to loading cargo, a number of checks are made to insure the suitability of the receiving airbase, airdrop zone, or opportune site. If no feasible link to the delivery site exists, the cargo is blocked from further progress.

Next, combat and attrition within GAMM were covered. Inflight survivability depends on three factors: originating airbase distance from the FLOT, receiving airbase distance from the FLOT, and aircraft type. GAMM currently does not model damage and attrition to aircraft while they are on the ground.

The three scenarios currently available for use with GAMM were next discussed. The Europe (EUR), Southwest Asia (SWA), and Central America (CA) scenarios can be used to assess tactical airlift system performance over a range of operating conditions. The SWA scenario was selected for this study because it is vastly different than the CA scenario which was used in an earlier analysis of airlifter characteristics. The two scenarios span a wide range of tactical airlift operating environments. As a result, the findings of the two

studies can be compared to make some generalizations about significant tactical airlifter characteristics.

Finally, to provide confidence that GAMM is a suitable analytical tool for exploring tactical airlift system performance, some of the more extensive verification and validation efforts carried out to date were covered. These efforts included verification of computer algorithms, conceptual model validation, operational validation, and data validation. The scope of the work done in the area of model verification and validation point to the fact that, "The model has been fully screened and is ready to be used in an analysis of FTAS conceptual aircraft" (6:125).

III. Response Surface Methodology

Introduction

A *response* can be defined as the output of a system or process, given a specific set of inputs. Experiments are often designed to quantify the relationship between a particular response and a group of factors (inputs) thought to affect it. In many cases, the goal of such an experiment is to identify the values of the inputs that yield the best value of the output. For example, the chemist might know that the purity of an end product of a chemical process is affected by the concentrations of several reagents in the solution and the reaction temperature. As a result, the chemist would try to quantify the relationship between inputs (concentrations of the reagents and reaction temperature) and the output (purity of the product). With this relationship quantified, the chemist might be able to determine the specific combination of reagent concentrations and temperature level that results in the purist product.

Response Surface Methodology (RSM) can be described as a scientific method of developing an *empirical model* for a response system when little is known about the *theoretical model* that governs that system. In the chemical reaction example given above, it might be known that the response $\mathbf{y}$ is given by $f(\boldsymbol{\Phi},\boldsymbol{\Theta})$ where the expected value of the response is

$$E(y) = \mu = f(\boldsymbol{\Phi},\boldsymbol{\Theta}) \qquad (2)$$

where $\boldsymbol{\Phi} = (\Phi_1,\Phi_2,\ldots,\Phi_k)$ is a set of input variables that measure initial concentrations of reactants and reaction temperature, and $\boldsymbol{\Theta} = (\Theta_1,\Theta_2,\ldots,\Theta_p)$ represents a set of physical parameters measuring such things as activation energies, diffusion coefficients, and thermal conductivities that define the functional relationship between the

inputs and response. In this case, there is complete knowledge of the relationship between the inputs, or factors, and the response. The response function $f(\Phi,\Theta)$ is the theoretical model for this process (1:11).

Often, however, the physical knowledge of the system is incomplete and an adequate theoretical model is not available. If it can be assumed that the theoretical model is somewhat smooth and continuous, which is generally the case, response surface methodology can be employed to approximate the theoretical model with an empirical model. When the *region of interest*, or the region defined by the range of input values, is relatively small, the empirical response surface can often be modeled as a low-order polynomial. The primary goals of RSM are to:

1. Design a series of experiments that will yield adequate and reliable measurements of the response of interest in a region of interest.
2. Analyze the results of those experiments to determine a mathematical model that best fits the data.
3. Find the settings of the input variables that produce the optimal response within the region of interest (5:3).

To achieve the goals listed above, response surface methodology employs a set of statistical and mathematical techniques. The techniques include:

1. <u>Designed Experiments</u> -- A process of experimentation in which purposeful changes are made in the input variables so that we may observe and identify the reasons for changes in the response (12:1).
2. <u>Regression Analysis</u>-- Statistical techniques used to model the response as a linear combination of various forms of the input variables and their interactions.

3. <u>Steepest Ascent</u> -- a gradient search technique for finding local maxima.

The first step in developing the empirical model is to choose some number of experiments to be performed at various values, or *levels*, of the factors. At each combination of factor levels, or *design point*, an observed value of the response is recorded. A regression analysis is then performed on the observed responses to arrive at the form of the empirical model, $g(\mathbf{\Phi},\mathbf{B})$ (4:4). The elements of the vector $\mathbf{B}$ are coefficients in the interpolation function which are related to, but distinct from, the $\mathbf{\Theta}$'s of the physical system. Again, the only assumption required for this analysis is that the underlying theoretical model, $f(\mathbf{\Phi},\mathbf{\Theta})$, is continuous over the region of interest.

The description of the RSM process given above indicates that it is actually a method for conducting scientific experimentation. Experimental investigation is not a haphazard or arbitrary process. Prior to conducting any experiment, the following questions should be considered:

1. Which input variables should be considered?
2. What is the appropriate response?
3. Should the input or output variables be transformed or used in their original form?
4. At which levels (values) of the input variables should experiments be run?
5. How complex does the model need to be?
6. How will qualitative factors be considered?
7. What experimental designs should be used (1:4-6)?

If some, or all, of the questions above are answered incorrectly initially, corrections can be made along the way. Response surface methodology is an iterative learning process. Experimental designs can be *augmented*, *transformations* can be changed, and *design regions* can be

modified. It is more important to rapidly converge to appropriate conclusions than to get it right the first time.

The remainder of this chapter will discuss designed experiments and regression analysis techniques used in a typical RSM analysis. Steepest ascent will not be covered since it was not used in this research.

Designed Experiments.

In the absence of a theoretical model, an experiment can be designed to facilitate the development of an empirical model. There are many experimental designs available, and the worth of the empirical model depends heavily on the design used. For example, *Box-Behnken* and *Central Composite* designs are very useful designs for estimating *second-order models* and should be considered any time quadratic terms are believed to be significant. On the other hand, *two-level factorial* designs are of great value for estimating *first-order models* and models containing only first-order and interaction terms (no quadratic, cubic, etc... terms). In addition to the degree of the model being estimated, some other important design characteristics to consider include:

- the ability to check the fit of the model to the response surface,
- the ability to estimate response transformations,
- the ability to augment to fit higher order models, and
- the number of runs required (1:502).

Factorial Experiments and the 2^k Fractional Design. A very common experimental design is the *factorial*. A *full* factorial design in *k* factors is obtained by choosing n_1 levels of factor 1, n_2 levels of factor 2,...,n_k levels of factor k. The n runs are selected such that all possible combinations of factor levels are covered in $n = n_1 \times n_2 \times \ldots \times n_k$ runs. Factorial designs include the following desirable properties:

1. They allow many simultaneous comparisons to be made.

2. They yield highly efficient parameter estimates. This means the estimates have variances that are as small, or nearly as small, as those that could be produced by any design occupying the same space.

3. They are computationally simple to analyze (1:106).

In a 2^k design, only two levels (high and low) for each factor are considered. The 2^k experiment is designed to adequately and efficiently measure the *main effects* and *interaction effects* of the factors on the response. A main effect in a designed experiment is defined as the change in response produced by a change in the level, or value, of that factor. Interaction effects occur when the difference in response between the levels of one factor is not the same at all levels of the other factors. Or, said differently, if there is no interaction, then the effect of a factor is the same regardless of the levels of the other factors. There are many good reasons for fitting a first-order model.

> First, when the region of interest is small in relation to the factor space, it is assumed that the response surface can be approximated reasonably well by the hyperplane defined by a first-order model. Also, when it is desirable to obtain *a first approximation of the surface* the simplest form of the model is fitted initially to assure that the time and cost of experimentation are held to a minimum. Still another reason for fitting a first-degree model is when the experimenter is *screening for the most important factors* (4:39).

Fractional Factorial Designs. To reiterate, a full factorial design includes all possible combinations of factor levels considered. Since each experiment can consume valuable time, money, or resources, implementing a full factorial design is often not feasible. There is a another good argument for not using a full factorial design. Consider the 2^k factorial arrangement

$$Y = \beta_0 + \sum_{i=0}^{k} \beta_i X_i + \sum\sum_{i<j}^{k} \beta_{ij} X_i X_j + \sum\sum_{i<j<l}^{k}\sum \beta_{ijl} X_i X_j X_l + \cdots + \beta_{12\ldots k} X_1 X_2 \ldots X_k + \epsilon \tag{3}$$

In general, a full 2^k design allows estimation of the mean plus 2^k-1 other terms. Only the first k + 1 terms of the model define the equation of a hyperplane. The remaining $2^k - (k + 1)$ terms consist of all the possible cross-products among the factors and measure the distortion of the hyperplane due to factor interactions (4:43). When the interaction effects are negligible, a reduction in the number of terms in the model can be accepted. Since fewer than 2^k terms are being estimated the number of design points can also be reduced. A 2^k *fractional replicate* is an experimental design that considers only a subset of the 2^k factor-level combinations. For example, a 1/2 fractional replicate, denoted by 2^{k-1}, contains only one-half the number of design points of the full 2^k arrangement.

Fractionating 2^k designs, while reducing the required number of runs, does not come without a cost. When only considering a fraction of all the possible factor level combinations, effects can become *aliased* with each other. When two or more effects are aliased, they cannot be measured separately and are said to be confounded (4:47). This problem can be minimized by wisely choosing the fraction to be considered. Since third-order and higher-order interaction effects are often found to be negligible, it would be desirable to design the experiment such that main effects and second-order interactions are only confounded with third-order and higher-order terms.

The resolution of a 2^k fractional factorial design defines the nature of its *alias structure*. If a 2^{k-p} design is of resolution **R**, then no **f**-factor effects will be confounded with any other effect containing less than **R-f** factors. Using the above definition, one can see that with a resolution V design no main effects (f=1) are aliased with any other main effects or with any two- or three-factor interactions, however two-factor (f=2) interactions are aliased with three-factor interactions or higher. Two-level, resolution V (2_V^{k-p}) designs are quite appropriate designs for fitting second-order models

when quadratic terms and three-factor interactions (and higher) are not significant.

Coded Variables. Since it is often cumbersome to work with the actual input values, the X's in Equation (2) are usually expressed as *coded variables*. The coding scheme most commonly used defines the coded variable, x_{ui}, in standardized form as

$$x_{ui} = \frac{X_{ui} - \overline{X}_i}{S_i} \qquad i = 1, 2, \dots, k \tag{4}$$

where

X_{ui} = the level of factor X_i at the u^{th} design point,

$\overline{X}_i$ = the mean of the X_{ui} values, and

S_i = some scale factor.

For example, consider the two-level factorial experiment where each of the k factors is set to either a high value (X_{HIGH}) or a low value (X_{LOW}). If the same number of observations are collected at each level, then $\overline{X}_i = (X_{LOW} + X_{HIGH})/2$ and $S_i = (X_{HIGH} - X_{LOW})/2$. In this case, the values of the coded variable x_{ui} is -1 when $X_{ui} = X_{LOW}$ and +1 when $X_{ui} = X_{HIGH}$. Note that the coded quantities x_{ui} are simply convenient linear transformations of the X_{ui}, and so expressions containing the x_{ui} can easily be rewritten in terms of the X_i.

Designs for Fitting Second-Order Models. If there is reason to expect significant quadratic effects, a design for fitting a second-degree model should be used. Consider the second-order model in k variables

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum\sum_{i<k}^{k} \beta_{ij} x_i x_j + \epsilon \qquad (5)$$

where the β_i are regression coefficients for the first-order terms, the β_{ii} are the coefficients for the pure quadratic terms, the β_{ij} are the coefficients for the two-factor interaction terms, and ε is the random error term. When estimating second-order models three or more factor levels must be considered. The 3^k factorial arrangement can be used to fit a second-order model, but as the number of factors increase the number of observations required ($3^5 = 243$ and $3^9 = 59{,}049$) can become impractical. Two arrangements for fitting second-order models that are much more efficient, in terms of the number of experimental runs required, are the *central composite design* and the *Box-Behnken design*.

<u>The Central Composite Design (CCD).</u> The CCD in k variables consists of:

1. The 2^k vertices of a k-dimensional "cube", or a 2^{k-m} fractional replicate of at least resolution V, where the factor levels are coded as in Equation (4) so that the design center is at $(x_1, x_2, \ldots, x_k) = (0, 0, \ldots, 0)$. These vertices can be thought of as the factorial portion of the design.
2. The $2k$ vertices $(\pm\alpha, 0, 0, \ldots, 0), (0, \pm\alpha, 0, \ldots, 0), \ldots, (0, 0, \ldots, \pm\alpha)$ of a k - dimensional "star".
3. The $n_0 \geq 1$ center point replicates (4:52).

The total number of experimental runs required is $n = 2^k + 2k + n_0$ or $n = 2^{k-m} + 2k + n_0$ when a fractional replicate is used. One obvious advantage to this design is that it can be formed by augmenting a two-level factorial design, leading to obvious efficiencies in sequential experimentation. If a model developed from a two-level factorial design proves to be inadequate, it is a simple matter to add the necessary star points and center point replications to the factorial design to create a CCD, which can be used to estimate a quadratic model.

Whether or not a CCD possesses the desirable properties of *orthogonality* and *rotatability* depends on the values selected for n_0 , the number of center point replications, and α, the distance the star points are from the design center. An orthogonal design allows the individual effects of the k variables to be assessed independently. A rotatable design gives equal precision for fitted responses at design points that are at equal distances from the design center. One disadvantage of CCD designs is that star points that allow orthogonality and rotatability might represent factor level combinations that are not feasible. Another disadvantage is that factors must be set to five levels as opposed to only three levels for many second-order designs.

Box-Behnken Designs. Box-Behnken designs are a subset of 3^k factorial designs and were designed specifically for fitting second-order models. "The designs are formed by combining 2^k factorial designs with incomplete block designs" (13:220) and require a relatively small number of experimental runs. The designs are not always orthogonal or rotatable, but are, at least, "nearly orthogonal" and "nearly rotatable". They require only three levels for each factor and do not contain any points at the extremes of the cubic region created by the two-level factorial. It is a very useful design when the extreme points of the cubic region represent factor level combinations that are prohibitively expensive or infeasible due to physical constraints (13:220). However, the lack of "corner points" in the cubic region precludes estimation of three factor interactions.

Regression Analysis

One of the goals of RSM is the development of an empirical model which adequately represents the theoretical model over a given region of interest. This section covers the techniques used in developing this empirical model. These techniques are designed with the following objectives in mind:

1. The selection of the *best model* among the set of plausible models.

2. The *efficient estimation* of the parameters (coefficients) in the selected model (4:11).

This discussion is limited to the development and testing of a first-order empirical model, but the results generalize to higher-order cases. Much of the information in this section is adapted from How to Apply Response Surface Methodology, by John A. Cornell (4).

Least Squares Estimation of an Empirical Model. Much of RSM analysis deals with general linear models. The polynomial representation of a first-order response surface can be written as a first degree model in *k* variables as follows:

$$Y_u = \beta_0 + \sum_{i=1}^{k} \beta_i x_{ui} + \epsilon_u; \quad u = 1,2, \ldots ,N; \quad i = 1,2, \ldots ,k. \tag{6}$$

where

N = the number of trials,

Y_u = the observed value of the response in the u^{th} design point,

x_{ui} = the value (or level) of the i^{th} controllable factor at the u^{th} design point,

β_0 and β_i = represent unknown parameters (coefficients) to be estimated, and

ε_u = the error made when observing Y_u .

The first step in fitting a first-order model to approximate the response surface is to collect data and estimate the $k + 1$ unknown coefficients, β_0 and β_i. After the coefficients are estimated, the estimates are then substituted into the equation yielding the estimated response equation, or empirical model:

$$\hat{Y} = b_0 + \sum_{i=1}^{k} b_i x_{ui} + e_u \qquad (7)$$

where

$\hat{Y}$ = the estimated value of the response, and

b_0 and b_i = the estimates for β_0 and β_i.

The *method of least squares* selects as the estimates b_0, b_1, b_2, ...b_k, for the unknown coefficients, β_0 ,β_1, β_2, ..., those values that minimize the quantity

$$R(\beta_0, \beta_1, \ldots) = \sum_{u=1}^{N} (Y_u - \beta_0 - \sum_{i=1}^{k} \beta_i x_{ui})^2. \qquad (8)$$

$R(\beta_0, \beta_1, \ldots, \beta_k)$ is the *sum of squares function,* which represents the sum of the squared *errors,* or residuals, between the experimental response (or observation) and the postulated model.

Matrix Form of Least Squares. In matrix notation, the first-degree model of Equation (7) over the $\boldsymbol{N}$ observations, can be written as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{\epsilon} \qquad (9)$$

where

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Nk} \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}$$

The sum of squares function, Equation (8), can be written as

$$R(\mathbf{b}) = \boldsymbol{Y}'\boldsymbol{Y} - 2(\boldsymbol{X}\boldsymbol{b})'\boldsymbol{Y} + \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}$$

Now setting $\frac{\partial}{\partial \mathbf{b}} R(\mathbf{b})=0$ in order to locate a stationary point gives

$$(\mathbf{X}'\mathbf{X})\mathbf{b}=\mathbf{X}'\mathbf{Y},$$

which is the matrix form of the normal equations. Solving these for **b** gives

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad (10)$$

The b_i's are the least square estimates for the β_i's in Equation (6). It is assumed here that the inverse of $\mathbf{X}'\mathbf{X}$ exists (i.e. $\mathbf{X}'\mathbf{X}$ is *non-singular*).

Assessing the Worth of a Regression Model. This section will be concerned with the first objective of linear regression stated earlier; selection of the *best model* among the set of plausible models. Ideally, a regression model will:

1. Be significant (as indicated by the *F-test*).
2. Exhibit a small error component (*Mean Square Error*).
3. Include a small number of parameters.
4. Account for almost all the variation in the response about its mean (as measured by R^2).
5. Provide an adequate fit of the surface over the region of interest.

The following sub-sections will cover some of the tests, techniques, and statistics used when determining model adequacy.

Analysis of Variance (ANOVA). ANOVA is a method of bookkeeping for the results of a linear regression analysis on a set of experimental data. The entries in an *ANOVA table* represent the sources that contribute to the total variation in the data values (4:17). The *total sum of squares* (SST) is the sum of the squared responses and the *corrected total sum of squares* (SSTc) is the total variation in the observed data values about their mean. SST and SSTc are defined by

$$SST = \sum_{u=1}^{N} Y_u^2 \qquad (11)$$

and

$$SSTc = \sum_{u=1}^{N} (Y_u - \overline{Y})^2. \qquad (12)$$

The *sum of squares due to regression* (SSR) is the sum of the squared predicted responses and *corrected sum of squares due to regression* (SSRc) is the sum of squared predicted responses corrected for the mean observed response. SSR and SSRc are given by

$$SSR = \sum_{u=1}^{N} \hat{Y}_u \qquad (13)$$

and

$$SSRc = \sum_{u=1}^{N} (\hat{Y}_u - \overline{Y})^2. \qquad (14)$$

The SSRc represents a measure of the difference between the value predicted by the fitted model and the overall average of the observed responses.

Finally, SSE is the sum of squares not accounted for by the fitted regression model. The formula for the SSE is

$$SSE = \sum_{u=1}^{N} (Y_u - \hat{Y}_u)^2 \qquad (15)$$

Table 2 is an example of an ANOVA table for corrected sums of squares. The *mean square* in each row is simply the associated sum of squares divided by its *degrees of freedom* (df). Note that the degrees of freedom and sums of squares are additive; the mean squares are not.

Table 2. The Analysis of Variance Table

Source	df	Sum of Squares	Mean Square
Due to Regression (fitted model)	$p-1$	SSRc	MSR = SSRc/(p-1)
Residual	N-p	SSE	MSE = SSE/(N-p)
Total	N-1	SSTc	

The F-Test for Model Significance. Once the ANOVA table is constructed, the usual *test of significance of the fitted regression equation* can be conducted. The null hypothesis of the test is H_0: All of the β_i, excluding β_0 if using the corrected sums of squares, are zero. The alternative hypothesis of the test is H_0: At least one of the β_i, excluding β_0 if using the corrected sums of squares, is not zero. The only additional assumption required for conducting the test is that the errors (ε_i) are normally distributed. If this assumption is satisfied, then the **F** statistic

$$F = \frac{\textbf{Mean Square Regression}}{\textbf{Mean Square Residual}} = \frac{SSR/(p-1)}{SSE/(N-p)} \qquad (16)$$

is computed for the corrected model. The computed **F** statistic is compared to the critical value $F_{(p-1,N-p,\alpha)}$, which is the 100α quantile of the **F** distribution with p-1 and N-p degrees of freedom, respectively. If the computed **F** statistic exceeds the critical value, the null hypothesis is rejected at the α level of significance, inferring that at least one of the coefficient estimates is not zero. Said differently,

the variation accounted for by the model is significantly greater than the unexplained information, and one or more of the coefficient estimates presumably conveys information about the response surface (4:18).

The Coefficient of Multiple Determination (R^2).

The *coefficient of multiple determination,*

$$R^2 = \frac{SSRc}{SSTc} = \frac{\sum_{u=1}^{N} (\hat{Y}_u - \bar{Y})^2}{\sum_{u=1}^{N} (Y_u - \bar{Y})^2}, \qquad (17)$$

is a measure of how well the observed model conforms to the observed data. It represents the proportion of total variation of the observed responses about the mean explained by the regression equation (4:18). In Equation (17), the numerator is a measure of the variability of the fitted responses about the mean response, while the denominator is a measure of the total variability of the actual responses about the mean response.

One limitation with R^2 becomes evident when considering the case of a constant response. In this instance, a model that fits the data perfectly will return a coefficient of determination equal to zero. To get a more complete indication of the quality of the fit, the quantity SSR/SST, an uncorrected R^2 (R^2_{unc}), can be considered in addition to R^2. In the constant response, R^2_{unc} would return a value of 1, which is what is expected.

Testing for Lack of Fit. Lack of fit of the first-order model, Equation (6), can result from curvature in the shape of the response surface due to excluded pure quadratic (or cubic) terms or from interaction effects between the experimental factors (4:20). The test

for lack of fit requires the following two experimental design conditions be satisfied:

1. The number of distinct design points, *n*, must exceed the number of fitted terms in the model ($n > k + 1$).
2. *Replicate observations* must be collected at one or more design points in order to estimate the error variance.

When these conditions are satisfied, the residual sum of squares (SSE) can be partitioned into two sources of variation. The first source is due to lack of fit of the model. Lack of fit occurs when the model is misspecified; that is, some of the important explanatory variables were omitted from the model. The other source is pure error variation. To partition the SSE into these sources, first the *pure error sum of squares* (SSPE) is calculated as the sum of squares due to the replicate observations. The *sum of squares due to lack of fit* (SSLF) is then obtained by subtracting the SSPE from the SSE (4:20). The SSPE is given by

$$\mathrm{SSPE} = \sum_{l=1}^{n} \sum_{u=1}^{r_l} (Y_{lu} - \bar{Y}_l)^2 \qquad (18)$$

where

n = the number of design points,

Y_{lu} = the u^{th} observation at the l^{th} design point,

r_l = the number of observations at the l^{th} design point, and

$\bar{Y}_l$ = the average of the Y_{lu} at the l^{th} design point.

The SSLF then equals

$$SSLF = SSE - SSPE = \sum_{l=1}^{n} r_l(\hat{Y}_l - \bar{Y}_l)^2 \qquad (19)$$

where

$\hat{Y}_l$ = predicted value of the response at the l^{th} design point produced by the fitted model.

The degrees of freedom associated with SSPE is

$$\sum_{l=1}^{n} (r_l - 1) = N-n \qquad (20)$$

where N is the total number of observations collected at the n design points. The degrees of freedom associated with SSLF is df(SSE) - df(SSPE) = (N-p) - (N-n) = $n-p$, where $p = k + 1$ is the number of terms in the fitted model.

The *test of adequacy of the fitted model* is

$$F = \frac{SSLF/(n-p)}{SSPE/(N-n)} \qquad (21)$$

The null hypothesis of adequacy of fit is rejected at the α level of significance if the calculated F statistic is larger than the critical value, $F_{(N-p,N-n,\alpha)}$. If the null hypothesis is rejected, the first-order model can be enhanced by adding cross-product terms or higher degree terms in the x_i s. Note, this may require augmenting the first-order design with additional design points. Before augmenting the design, it might be wise to try a response or input data transformation to improve the fit. If the null hypothesis is not rejected, it can be inferred that the response surface is a plane. In this case the SSE is used as an estimate for σ^2 and is also used in the F-test for significance (4:21).

<u>Single Degree of Freedom Test for Curvature.</u> If there is significant lack of fit of the first-order, model an additional test can

be performed to determine if the lack of fit (SSLF) is due to the presence of *pure quadratic terms*. When replications are made at the center of the design region, the sum of squares due to lack of fit can be further partitioned to test for curvature due to quadratic terms. The *single degree of freedom test for curvature* compares the average response at the *corner* points of a first-order design with the average response at the center point of the design. In particular, if there are n_1 observed responses at the corner points of a first-order design and $n_0 > 1$ observed replications at the design center point then define

$$\overline{Y}_{n_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i \text{ as the average response at the corner points}$$

and

$$\overline{Y}_{n_0} = \frac{1}{n_0} \sum_{i=1}^{n_0} Y_i \text{ as the average response at the center point.}$$

A significant difference between the average response at the corner points and the average response at the center point is an indication that pure quadratic terms are important.

To further partition the SSLF the *sum of squares due to pure quadratic terms* (SSPQ) is first computed via

$$\mathrm{SSPQ} = \frac{n_0 n_1 (\overline{Y}_{n_1} - \overline{Y}_{n_0})^2}{n_0 + n_1} \qquad (22)$$

The SSPQ has a single degree of freedom associated with it. The portion of the SSLF due to other than quadratic terms is simply the difference between the SSLF and SSPQ.

The single degree of freedom test for curvature is performed by calculating

$$F = \frac{SSPQ}{SSPE/N-n} \qquad (23)$$

where N and n are defined as in Equation (21). If the calculated F statistic in Equation (23) is greater than the critical value $F_{(1, N-n, \alpha)}$, then curvature in the surface due to quadratic terms is suspected. If curvature is suspected and a better fit to the response surface is desired, the next step is to design and conduct an experiment for estimating a second-order model.

Analysis of Residuals. Studying the residuals $Y_i - \hat{Y}_i$ can provide considerable information relating to the adequacy of the fitted model. If assumptions concerning the adequacy of the model are true, the residuals should be normally distributed with a zero mean and constant variance. Residuals that do not conform to these assumptions question the adequacy of the fitted model. "Therefore, as an *automatic preliminary* to further statistical analysis, a plot of residuals against predicted responses should always be constructed and studied" (2:183).

A plot of residuals against the predicted responses can indicate nonlinearity or non-constant variance. For example, if the fitted model is appropriate, the residuals should be unrelated to the value of the response. If the variance of the residuals increases as the value of the response increases, the residual plot would take on a *funnel-like* appearance, indicating a non-constant variance. Often this trend can be corrected by an appropriate response transformation. Nonlinearity might be indicated by a curvilinear relationship between residuals and predicted responses (2:220). Here the residuals may tend to be positive for low values of the predicted response, become negative for intermediate values, and be positive again for high values. Again, this trend may be corrected by a transformation or it may point to the need for higher order terms in the model.

Another tool that can be used to verify the normality assumption is the rankit plot. The ith-rankit is defined as the expected value of the i-th *order statistic* for the sample, assuming the sample was from a normal distribution. The order statistics of a sample are the sample values reordered by their rank. If the residuals are normally distributed, a plot of their rankits against their order statistics should be nearly linear.

Reduction of the Number of Independent Variables (15:435-455). There are a number of reasons to reduce the number of independent variables to be used in the final model.

> A regression model with a large number of independent variables is difficult to maintain. Further, regression models with a limited number of independent variables are easier to work with and understand. Finally, the presence of many highly intercorrelated independent variables may add little to the predictive value of the model while substantially increasing the sample variation of the regression coefficients, detracting from the model's descriptive abilities, and increasing the problem of roundoff errors. (15:436).

One commonly used computerized method for reducing the number of independent variables is the *forward stepwise regression procedure.* Using the search algorithm outlined in Appendix C, the forward stepwise regression procedure develops a sequence of regression equations. An independent variable is either added or deleted at each step. Essentially, the search method enters the variable most highly correlated with the response, given the variables already in the model. It then examines whether any of the variables already in the model should be deleted, given the new variable. This process continues until none of the remaining variables (i.e. all the variables not already in the model) exceed the criterion for entering the model, usually defined in terms of a *partial* F statistic. Note, that this procedure allows a variable brought into the model at an earlier stage to be subsequently dropped if it is no longer helpful in the presence of variables added at later stages.

Summary

This chapter introduced response system methodology as a set of techniques used in experimentation and empirical model building. Two of these techniques, designed experiments and regression analysis, were covered in detail. The section on designed experiments showed how the worth of an empirical model depends heavily on the experimental design used. Two-level factorial arrangements were introduced as common designs for fitting models that contain only main effect and interaction terms. Box-Behnken and central composite designs were described as common designs for fitting second-order models. The regression analysis section covered least squares estimation and a number of techniques used to determine model adequacy. The chapter ended with a description of the stepwise forward selection process, a common technique for removing insignificant factors from the model.

The techniques covered in this chapter, along with the Generalized Air Mobility Model (GAMM), are used in this thesis research to analyze the impact of varying airlifter characteristics on airlift system performance. Using a sequential experimentation process, this study attempts to develop adequate, parsimonious models that relate airlifter characteristics to airlift system effectiveness measures provided by GAMM. These models should demonstrate which airlifter characteristics most impact airlift system performance. Such models could be used for a number of different purposes. For instance, they could be used to determine which areas to concentrate on when designing the next generation tactical airlifter. Also, if considered accurate enough, they could be used as metamodels to measure airlift system performance instead of running GAMM. At the very least these models should provide some very useful insights into the relationship between tactical airlifter characteristics and tactical airlift system performance.

IV. Experimental Approach

Introduction

This research uses GAMM as the basis of an experiment designed to identify airlifter characteristics which have the greatest effect on tactical airlift system performance. This chapter covers the measures of effectiveness considered appropriate for this research, experimental design selection, selection and grouping of experimental variables, and suitable values for experimental variables. It ends with a description of the stepwise regression approach used in this study.

Defining Measures of Effectiveness

"In many cases the choice of the proper *measure of effectiveness* of the operation is the all-important decision to make in the analysis" (14:4). Measures of effectiveness (MOEs) historically used for theater airlift systems include: rate of cargo movement, average aircraft flying time per day (utilization (UTE) rate), departure reliability, and the ratio of hours flown to hours scheduled to be flown. These measures of tactical airlift capability are primarily efficiency measures that do not take into account what really matters: The ability of tactical airlift to meet the user's needs. "The fact that tactical requirements are often determined by the user as a result of changing combat conditions makes response to these requests more important than the need to efficiently use the aircraft" (3:5).

Three of the four MOEs used in this research measure how effectively the system meets demand: *Ratio on Time* (ROT), *Ratio Delivered* (RD), and *Ratio of Critical Cargo Delivered* (RCC). RD and ROT refer to the total cargo demand which (1) was delivered, and (2) was delivered on time, respectively. RCC, for the purposes of this experiment, refers to that percentage of total *priority 1* and *priority 2* cargo delivered. In GAMM, priority 1 or 2 are defined as critical moves

and emergency resupply that, by their very nature, are time-sensitive from an operational perspective. A priority 1 or 2 requirement is deleted if it is more than a day late (6:76). Critical cargo is that cargo required by the army to prosecute the war. Any tactical airlift system not capable of meeting critical cargo movement demands satisfactorily would be considered highly deficient (22).

The fourth MOE used in this experiment, *Millions of Ton Miles Per Day* (TMPD), is a common measure of strategic airlift effectiveness. Although GAMM models a theater tactical airlift system, the Southwest Asia Scenario covers a large geographical region. As such, this strategic airlift MOE and the tactical airlift MOEs discussed above might identify the same set of airlifter characteristics as significantly impacting tactical airlift performance.

Because GAMM is a theater airlift system model it does not provide a value for TMPD, or all the information required to calculate it exactly. If there were *k* productive, or cargo carrying, sorties during the 30 day war, then millions of ton miles per day could be determined by

$$\text{TMPD} = \frac{1}{(30)(1{,}000{,}000)}\sum_{i=1}^{k} d_i t_i \qquad (24)$$

where d_i is the distance traveled by sortie i and t_i is the cargo, in tons, carried by sortie i. GAMM does not provide the distance traveled (d_i) and cargo carried (t_i) for each productive sortie. The model does provide the total tons delivered (TTD) , total number of productive sorties (TPS), and total productive flight hours (TPFH) for the 30 day simulation. The average distance traveled per productive sortie (ADT) is then (TPFH x cruise speed)/TPS. Assuming a constant distance travelled for each sortie results in

$$\text{TMPD} = \frac{\text{ADT}}{(30)\times(1{,}000{,}000)} \sum_{i=1}^{k} t_i, \qquad (25)$$

which reduces to

$$\text{TMPD} = \frac{(\text{ADT}) \times (\text{TTD})}{(30)\times(1{,}000{,}000)}. \qquad (26)$$

The assumption that the distance traveled per productive sortie is a constant, equal to the average distance traveled, seems reasonable for this scenario. Many of the originating airbases are in Saudi Arabia or other locations on the opposite coast of the Persian Gulf from Iran. This means a large portion of the productive sortie is spent crossing the gulf and this portion should be roughly equal for all productive sorties. Distances traveled once over Iran will then be small in comparison, therefore there should not be a large variation in the distance traveled by productive sorties.

As described, the value used for TMPD is an estimate. Although it would be desirable to be able to calculate TMPD exactly, it should be sufficiently accurate determine if a strategic airlift MOE indicates the same significant airlifter characteristics as the tactical airlift measures considered.

Selection of an Experimental Design

A 2_V^{8-2} design is implemented in this experiment. The design contains 64 design points plus four center point replications used for lack of fit tests. The primary disadvantage of this design is its inability to fit a quadratic model. Main effects and two-way interactions can be estimated, but any effect due to quadratic terms cannot. The decision to accept this restriction was based on the following considerations:

1. Metamodels developed by Pappas, using a similar experimental design, were highly significant and involved only main effects and two-way interactions (no quadratic terms) (16:80).
2. The primary goal of this experiment is to identify significant characteristics, through an adequate least squares fit of the data. If the empirical models developed from this experiment are reasonably good, some lack of fit due to the absence of quadratic terms might be considered acceptable.

Should the models resulting from the 2_V^{8-2} design prove to be inadequate, data can be generated at the necessary design points to support a second-order model. While a central composite design (CCD) is typically used to supplement a fractional factorial design, introducing two additional factor levels renders the CCD inappropriate for this experiment. The new lower level would unfortunately result in an aircraft too small to effectively support the Southwest Asia theater airlift mission. In lieu of a CCD, a Box-Behnken design that requires only three levels to estimate main effects, 2-factor interactions, and quadratic terms would be more appropriate.

Selection and Grouping of Variables

A two-level full factorial experiment assessing all the airlifter characteristics available with GAMM would require 2^{69} design points. Since this is not possible, two techniques were used to logically reduce the number of variables.

First, a number of parameters remained fixed. Parameters held constant included Taxi, Takeoff, and Landing Time; Reserve Fuel; Mean Time to Service; and Airfield Temperatures at Sea Level and at 5000 ft.

After removing these parameters from consideration, 39 factors remained. These remaining factors were considered significant and were grouped into functional sets to make the number of variables in the

experiment more manageable. Groupings were formed by combining parameters that were closely related in aircraft function or purpose. Individual characteristics that were highly correlated were also grouped together so that all group combinations, or design points, would represent feasible aircraft designs.

> For example, cargo cabin size is related to the parameters for cargo bay width, height and length, and the size of the cabin directly affects the size of the aircraft (Aircraft Spot Factor). Generally, as aircraft increase in size, all cabin dimensions will increase and the aircraft's Maximum Cabin Payload will also tend to increase. (16:35)

The four functional sets used in this experiment were: *Field Performance* (F), *Aircraft Size* (S), *Inflight Performance* (I), and *Ground Flotation/Wheel Loading* (G). The parameters included in each of these sets are listed in Table 3. These are the same groupings used in the Pappas experiment, with the following exceptions:

1. The Inflight Performance group now only includes *cruise speed*. *Maximum ferry fuel* and *cruise fuel* were moved to the Aircraft Size (Pappas's *Cargo Cabin*) group because their values are determined by the aircraft's size. Also, *takeoff/landing fuel bias* was removed from the Inflight Performance group, and the experiment, due to its perceived insignificance (22).
2. Pappas's Servicing and Aircraft Loading/Unloading group was removed entirely. It did not appear in any of the parsimonious models developed in the Pappas experiment. In other words, this group was not significant in the Central American scenario, where a relatively large percentage of time is spent on the ground due to short flight times. It was assumed that in an environment such as Southwest Asia, where a relatively small percentage of time is spent on the ground due to longer flight times, the effect of this group would also

Table 3. GAMM Airlifter Functional Groupings

Field Performance (F):
for sea level and at 5000 ft,
for hot and cold conditions.
CTOL TO AT MAX USEFUL LOAD (FT)
CTOL LD AT MAX USEFUL LOAD (FT)
CTOL TO AT MID USEFUL LOAD (FT)
CTOL LD AT MID USEFUL LOAD (FT)
CTOL TO AT ZERO USEFUL LOAD (FT)
CTOL LD AT ZERO USEFUL LOAD (FT)
Aircraft Size (S)
CARGO BAY WIDTH (INCH)
CARGO BAY HEIGHT (INCH)
CARGO BAY LENGTH (INCH)
CARGO BAY DOOR WIDTH (INCH)
CARGO BAY DOOR HEIGHT (INCH)
CTOL MAX USEFUL LOAD (LBS)
CTOL MID USEFUL LOAD (LBS)
MAXIMUM CABIN PAYLOAD (LBS)
AIRCRAFT SPOT FACTOR (NO)
CARGO THRESHOLD FOR RELOCATION (LBS)
MAXIMUM FERRY FUEL (LBS)
CRUISE FUEL BURN RATE (LBS/HR)
Inflight Performance (I)
CRUISE SPEED
GROUND FLOTATION/WHEEL LOADING (G)
LCN - MAX USEFUL LOAD (NO)
LCN - AT ZERO USEFUL LOAD (NO)

be insignificant. Finally, earlier studies by the Aeronautical Systems Center indicated that ground servicing times have little influence on airlift system performance (23).

Each group is considered as a single aggregate "factor", with all the individual factors inside a group set to their "high" or "low"

levels together. Effects are then determined based on the value of the response at the high and low levels of the aggregated factors.

Group screening does not come without its costs. Grouping the parameters confounds any analysis concerning individual parameters within a group, because the effect of any single factor will be combined with the effects of the other parameters within its group (11:678). This limitation is acceptable since the purpose of the research is to identify airlifter characteristics, which are represented by the aggregate factors, that significantly influence airlift system performance.

Characteristics considered in addition to those studied by Pappas include *Survivability* (P), *Reliability* (R), and *Maintainability* (M). Increasing (decreasing) any of these airlifter characteristics will make the airlifter more (less) available. If the airlifter fleet is working at or near capacity, increasing (decreasing) airlifter availability will increase (decrease) airlifter fleet productivity (24:542). Survivability, for the purposes of this research, is a probability as defined in Chapter II. Reliability is defined as the *Mean Time* (in hours) *Between Critical Failures* (MTBFc). A "critical failure" is a "grounding failure" since it requires repair prior to the next flight. MTBFc, a GAMM airlifter characteristic, is derived from failure rates of items on the airlifter's *Minimum Essential Systems List* (24:542). Finally, maintainability is defined as the *Mean Time To Repair* (MTTR), or the time required to effect a repair. Within GAMM, repair time for each failure is assumed to follow a lognormal distribution and the following assumptions are made in regard to MTTR:

1. If there is more than one failure, all the repairs are made simultaneously within the time of the longest repair.

2. If the airlifter is at a main operating base (MOB) there is an infinite supply of maintenance crews and spare parts and the repair action begins immediately.
3. An aircraft which fails at a facility other than a MOB is grounded until the maintenance support arrives. This *logistical pipeline delay* is accounted for by adding a bias time to the repair time. GAMM assumes a six hour logistical pipeline delay.(543:2).

Appropriate Factor Levels. In a two-level factorial design, the high and low values of the variables must be far enough apart to identify major trends in the response, such as the slope, but they cannot be so far apart that major features of the response, such as inflection points, are ignored (17:1142). Variable values in this study were selected around the C-130H baseline, so that significant findings could be directly related to the C-130H's capabilities.

For the this experiment parameter values were derived by multiplying the C-130H's characteristics by 4/3 and 2/3, as was done in the Pappas experiment. For the Aircraft Size (C) group the smaller cabin was then made slightly wider and higher to allow transportation of standardized packages within GAMM. Most standardized packages within GAMM have a base of 104 x 84 inches and a height of 96 inches. To ensure that the C-130H was still in the center of the experiment's design region, the width and height dimensions for the larger aircraft were slightly decreased. For most parameters, the smaller values were half the larger values.

Because this experiment examines the effect of functional groups of parameters, high and low values were required for each aggregate factor. The individual variables included in a functional group were simultaneously set to their high or low levels based on the level assigned to their functional group.

One very serious pitfall exists with the group screening approach applied to the aggregate factors. A particular group may in fact contain several important individual factors, but the group as a whole could appear insignificant if some of its member effects unexpectedly have an opposite impact and cancel each other out (11:678). This problem can be avoided by defining the "high" and "low" values of each individual factor based on its expected effect on the response. For example, the high level of an individual factor would correspond to that level which is expected to increase the value of the response.

In this experiment high and low levels of each individual factor were determined based on their expected effect on Ratio Delivered (RD). Variables were defined to be at their high level if they were expected to increase the RD. Otherwise, they were defined as low level. As a result, high and low levels do not always correspond to the respective larger and smaller values for the variables.

> For Field Performance (F), the highest value of Ratio Delivered is expected when aircraft can make use of shorter fields. The high level of this variable corresponds to the smaller values of takeoff and landing distances. For the Ground Flotation/Wheel Loading (G) variable, increased throughput is expected when the value of the parameter is smaller. (16:39)

While it would be desirable to define the levels of each factor within a group based on their expected effect, it may not always be possible in practice due to the interdependence among the factors within a group. Said differently, one factor being at a high level within a particular group may force another factor within the same group to be at its low level, due to some type of physical relationship or other interdependence. The Aircraft Size (N) group offers an example of individual effects within a group offsetting each other. Cruise Fuel has a value of 7067 lbs/hr at its high level and 3533 lbs/hr at its low level. A fuel flow of 7067 lbs/hr will actually decrease throughput compared to a fuel flow of 3533 lbs/hr. This anomaly cannot be avoided though, because the higher fuel flow is the result of a balanced

engineering design. A larger aircraft that carries more fuel will require a larger engine that burns fuel at a faster rate.

Although it is desirable to define factor levels within a group so that they do not offset each other, it is required that the factor levels are defined so that they are feasible. A larger, faster aircraft with a lower fuel flow is infeasible and any experiment that analyzed such a design would provide misleading results. Factor levels for this experiment have been selected so that no "absurd" designs are considered.

Within the Southwest Asia scenario (SWA), airbase probability of survival factors change on a weekly basis. The position of the Forward Line of Troops (FLOT) is updated at the end of each week and new airbase survivability factors are calculated as illustrated in Table 1 of Chapter II. The third week of the war includes the most intense fighting and, hence, the highest threat flying environment. To model the low level of Survivability (P), the model was run for the entire 30 day war using the airbase survivability factors from week three for a C-130. In other words, the low level of survivability represents 30 days of C-130 operations in a "high threat" environment. The probability of survival was set to one (impervious to threat) to model the high level of Survivability (P).

The goal for the next generation tactical airlifter is to double the MTBFc and halve the MTTR of a C-130E (24:543). For this research reliability and maintainability levels were set to encompass those values required by a follow on tactical airlifter as well as those of the C-130H. The Acquisition Logistics office recommended a C-130H MTBFc of 2.34 hours (18). Since C-130H and C-130E maintainability data are confounded, the Acquisition Logistics office advised a MTTR of 2.4 hours for a generic C-130 (18). Doubling these values, the next generation tactical airlifter requires a MTBFc of 4.68 hours and a MTTR of 1.2

hours. Varying MTBFc from 2 to 6 hours and MTTR from .6 to 3.0 hours will include the actual values for a C-130 and those required for the next generation airlifter (24:543).

Fleet Size (N) was the one operational consideration that was included in this experiment. Determining appropriate high and low levels for N involved making a number of test runs with GAMM. Varying Fleet Size (N) from 20 aircraft to 120 aircraft, while holding all other factors constant, indicated Fleet Size (N) had very little effect above 80 aircraft. As Figure 3 illustrates, increasing from 80 to 120 aircraft only increased Ratio on Time (ROT) from .40 to .41. On the other hand, there was a significant increase in ROT when going from 20 to 60 aircraft. In order to study the effect of N in a region where it appeared to have a significant impact, its low value was set at 20 and its high value at 60. It should be noted that there were similar results for Ratio Delivered, Ratio of Critical Cargo Delivered, and Millions of Ton Miles Per Day.

Appendix D summarizes the factor level settings used in this experiment.

Stepwise Regression

The observed responses and associated factor level settings are used to perform a multivariate regression analysis. A personal computer statistics package, *Statistix 4.0,* is used to perform a forward stepwise regression for each MOE. The following strategy is implemented to develop models to support the research objective:

1. First, a forward stepwise regression is performed, forcing all main effects and two-way interactions in the model. This is accomplished by setting the critical **F** value for entering and leaving equal to zero.
2. Next, the R^2 and MSE for each step are considered. Starting with the last variable entered, variables are removed until such

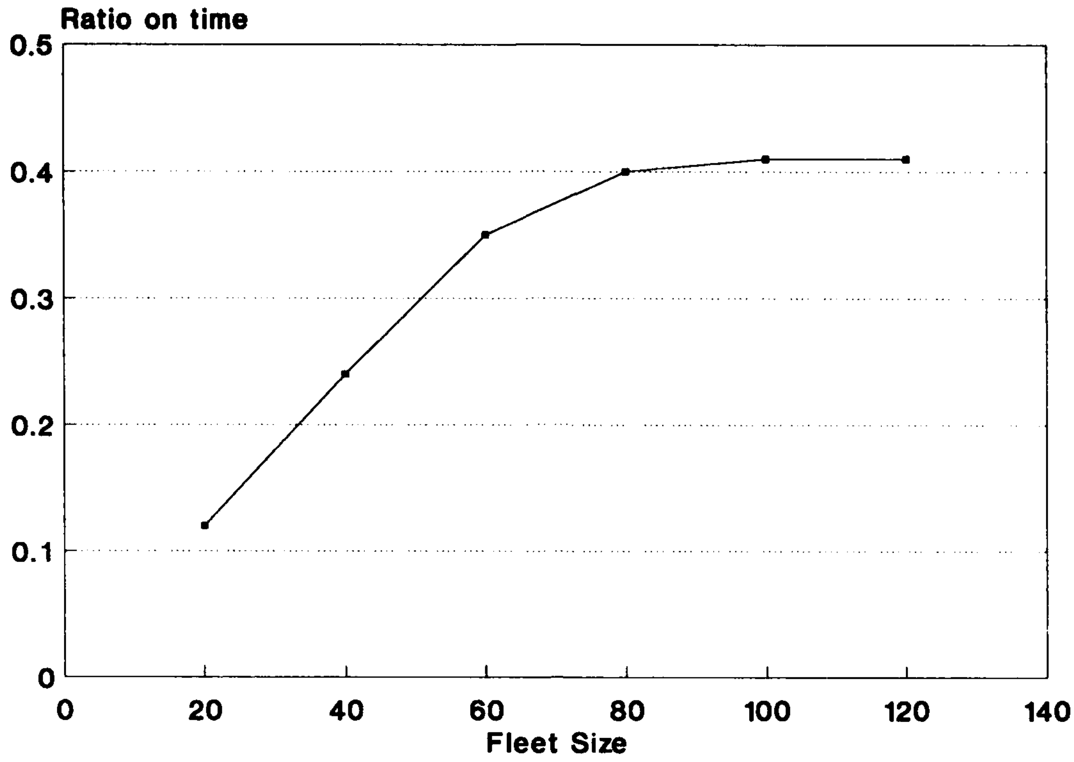


Figure 3. Range of Values for Fleet Size (N)

removal causes either a significant decrease in R^2 or a significant increase in MSE.

3. A minimum of 10 variables are retained so that a number of different possible models can be compared.

Plots of residuals against predicted values are analyzed to evaluate departures from standard assumptions.

The resulting model describes how departures from the baseline C-130H configuration affect the performance of the theater airlift system. The insignificant terms that do not appear in the model reflect aircraft characteristics that have little influence on system-level performance when varied about the baseline C-130H level. It should be emphasized, however, that if the values of the insignificant terms vary well beyond those examined in

this experiment, they could possibly have a major impact on system-level performance.

Summary

Four measures of effectiveness are considered in this research. Ratio Delivered (RD), Ratio on Time (ROT), and Ratio of Critical Cargo Delivered (RCC) are used to measure how well the system meets demand. The fourth measure, Millions of Ton Miles Per Day (TMPD), is used to determine if a traditional strategic airlift MOE indicates the same significant airlifter characteristics as the tactical MOEs when applied to a relatively large tactical theater, such as Southwest Asia.

A 2_V^{8-2} design is selected for this experiment. Forty-two aircraft parameters are grouped into four aggregate factors: Field Performance (F), Aircraft Size (S), Ground Flotation/Wheel Loading (G), and Inflight Performance (I). The other factors considered are Fleet Size (N), Maintainability (M), Survivability (P), and Reliability (R). A forward stepwise regression performed on the experimental results yields a least squares model that indicates which tactical airlift characteristics most influence tactical airlift system performance.

V. Experimental Results and Analysis

The 2_V^{8-2} Experiment

Factor level settings and response values for this experiment are listed in Appendix E. Center point replications were included so that lack of fit tests could be performed.

Models including main effects and two-factor interactions were developed with a stepwise regression procedure. Analysis of plots of residuals against predicted values revealed cases of non-linearity and non-constant variance. Some residual plots were curve shaped, suggesting the need for additional cross-product or higher-order terms in the model. Others had a funnel-like shape, implying a non-homogeneous error variance. In both cases, response transformations were used to correct the problem. ANOVA tables and residual plots of the transformed models were then compared to those from the original data. If the model developed from the transformed results proved to be better, it was used.

In some cases a square root (SQRT) transformation worked best, while in others a natural logarithm (LN) transformation was more appropriate. A transformed response is denoted by the abbreviation for the transformation immediately followed by the abbreviation for that measure of effectiveness. For example, SQRTRD represents the square root (SQRT) of Ratio Delivered (RD).

The models are provided in Figure 4. Adequacy measures for the models are summarized in Table 4. Residual plots indicated that residuals were randomly scattered about zero and rankits plots were nearly linear. A lack of fit test was conducted for each model. If this test was significant, a single degree of freedom test for curvature

Square Root of Ratio Delivered Model (SQRTRD)

$$\hat{Y} = .61318 + (.09188)N + (.07622)S + (.05175)P + (.03314)G + (.03293)I + (.03050)FG + (.02498)NF + (.01866)M$$

Natural Logarithm of Ratio on Time Model (LNROT)

$$\hat{Y} = -2.03878 + (.49098)N + (.39902)S + (.22254)I + (.17965)P - (.14683)F + (.13423)G + (.12143)FG + (.11978)FP + (11743)M$$

Square Root of Ratio of Critical Cargo Delivered Model (SQRTRCC)

$$\hat{Y} = .58553 + (.10118)S + (.08304)N + (.07391)P + (.04073)F + (.03744)I + (.03709)NF + (.03709)G + (.03627)FG$$

Square Root of Ton Miles Per Day Model (SQRTTMPD)

$$\hat{Y} = .49756 + (.07211)N + (.05960)S + (.03658)P + (.02835)I - (.02778)F + (.02531)G + (.02350)FG + (.01700)NF + (.01417)M$$

Legend:
N = Fleet Size
S = Aircraft Size
I = Inflight Performance
P = Survivability
F = Field Performance
G = Ground/Flotation
M = Maintainability (MTTR)

Figure 4. Regression Models for Two-Level Factorial Experiment

Table 4. Stepwise Regression for Two-Level Factorial Experiment

Eqn No and Added Factor	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
SQRTRD	.9200	.9957	.00198	84.84	.1131	.0490
LNROT	.9412	.9930	.03889	103.21	-.1165	.0354
SQRTRCC	.9114	.9927	.00313	75.90	.2631	.0715
SQRTTMPD	.9300	.9964	.00110	85.66	.0673	.0289

$\varepsilon_i = Y_i - \hat{Y}_i$ (Max ε_i is in terms of magnitude)
ε_i are in terms of the original data.

was performed. The results of these tests are summarized in Table 5 below.

Table 5. Summary of Lack of fit and Curvature Tests

Response	Lack of fit F value	Single-DF F Value	P
SQRTRD	80.23	326.3	9
LNROT	55.0	591.9	10
SQRTRCC	38.6	19.62	9
SQRTTMPD	139.8	709.2	10

α = .10
N = number of runs = 68
n = number of design points = 65
n_0 = number of center point replications = 4
p = number of terms in the model
Lack of fit Fc = $F_{(\alpha,n-p,N-n)}$ = 5.15
Single-DF Fc = $F_{(\alpha,1,n_0)}$ = 5.54

As Table 5 indicates, all models demonstrate significant lack of fit. Further, the single degree of freedom test for curvature for each MOE indicates the presence of pure quadratic terms. To determine which quadratic terms were affecting the response, a second-order experimental design capable of estimating the quadratic coefficients is required. The Box-Behnken design (BBD) was selected over the central composite design (CCD) because the BBD is "near-orthogonal" and "near-rotatable" while using the same high and low values used in the factorial arrangement. Achieving orthogonality and rotatability with a CCD would require factor level settings that are not feasible from an aircraft design perspective.

There was one additional important result from this experiment. As Figure 4 indicates, Reliability (R) is not present as a main effect or interaction effect in any of the models. Since R was insignificant in this experiment it is not considered in the Box-Behnken experiment. An earlier study carried out by the Aeronautical Systems Center (ASC)

also found Reliability (R), when varied around a baseline C-130, to have a relatively insignificant impact on airlift system performance (24:548).

Finally, there was not much variation in the responses for the four center point replications. Ratio of Critical Cargo Delivered (RCC) exhibited the widest range of values, only varying between .3640 and .3793. However, upgrading to the Box-Behnken experimental design represents a further refinement of the analysis. To reduce the variance, five replications are made at each design point in the Box-Behnken experiment.

The Box-Behnken Experiment

The Box-Behnken design and response values for this experiment are listed in Appendix F. Response values are actually the averages for the five replications. A stepwise regression was carried out for each MOE. After making either a square root or natural logarithm transformation, residuals were reasonably well scattered about zero and rankits plots were nearly linear. A minimum of ten terms plus the constant were retained for each MOE so that a number of different models could be compared. Correlation matrices for each MOE are shown in Appendix H. There were no significant correlations between independent variables.

The stepwise regression performed on each MOE will be covered after a discussion on the interactions between Survivability (P), Field Performance (F), and Ground Flotation (G).

Survivability (P), Field Performance (F), and Ground Flotation(G)

Field Performance (F) had a negative coefficient in the models developed for the natural logarithm of Ratio on Time (LNROT). A negative coefficient implies that the ability to operate into shorter strips is counter productive. This seemingly anomalous result can best be explained by considering the significant interaction terms involving

Field Performance (F). In the stepwise regression for LNROT, Field Performance (F) had a significant interaction with both Survivability (P) and Ground Flotation (G). The interaction plot for Field Performance (F) and Survivability (P) is shown in Figure 5. The top line represents a no threat (P+) environment while the bottom line represents the high threat (P-) environment used in the simulation. Ground Flotation (G) was at its middle level (0) for all design points used to compute the FP interaction. As expected, when operating in a no threat environment an increase in Field Performance (F) results in a slight increase in ROT. However, when operating in a high threat environment an increase in F results in a pronounced decrease in ROT. This decrease in ROT can be explained in terms of attrition. For the high threat, low field performance combination (P-,F-) an average of 4 aircraft are lost. For the high threat, high field performance case (P-,F+), an average of 14.5 aircraft are lost. Since Fleet Size (N) was at its middle value of 40 the latter case represents a 36% attrition rate over the course of the 30 day war. Knowing that Fleet Size (N) was significant, its obvious that a 36% attrition rate will significantly reduce productivity.

There is higher attrition in the high threat, high field performance case (P-,F+) because with improved field performance the airlifters operate into more of the short fields which tend to be near the threat, thus exposing them to the threat with increased frequency. In fact, there was an average of 458 flights in the vicinity of the threat when the aircraft were short field capable. This is almost four times the 117 average number of flights near the threat when the aircraft were not capable of short field operations .

The fields are typically shorter near the threat because within GAMM the intensity of runway attacks is a function of the distance from the FLOT. The closer the field is to the FLOT, the greater the

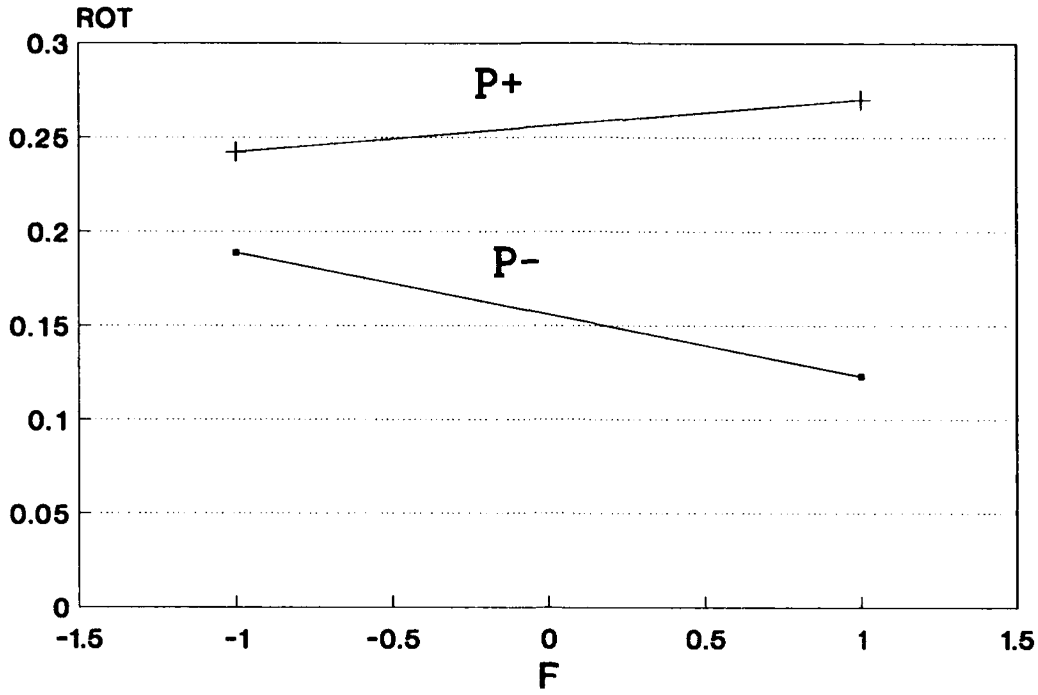


Figure 5. FP Interaction Plot

intensity of the attack. GAMM then uses attack intensity to determine the usable runway length after the bombing. The net effect; the closer a given runway is to the FLOT the less usable runway length there will be remaining after an attack.

The interaction plot for Field Performance (F) and Ground Flotation (G) is shown in Figure 6. Survivability (P) was at its middle level (0) for design points used to compute the FG interaction. The top line (G+) represents an aircraft that can operate into unprepared surfaces while the bottom line (G-) represents an aircraft that requires a hard landing surface. Improving Field Performance (F) when Ground Flotation (G) is at its high level results in a slight increase in ROT, although attrition increased in this case from 1.5 to 9 aircraft, or from 4% to 23%. There was an average of 106 flights near the threat

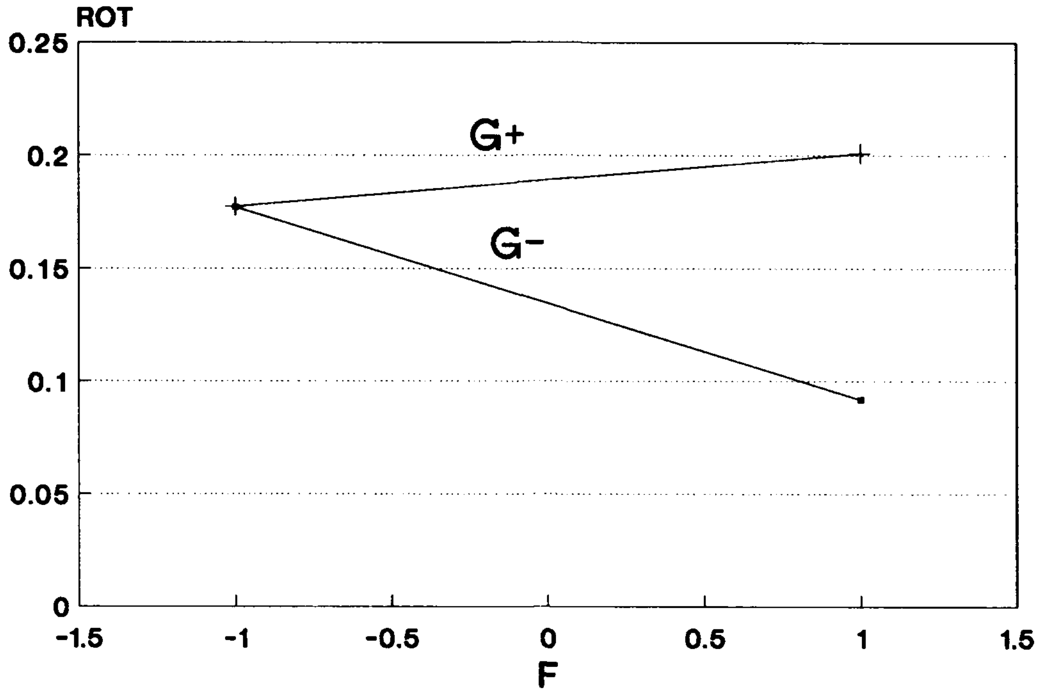


Figure 6. FG Interaction Plot

when the aircraft were limited to long, unprepared runways (F-,G+) operations and 519 flights near the threat when short, unprepared runway (F+,G+) operations were possible. Thus, a 5 fold increase in the number of flights near the threat caused a 6 fold increase in attrition, a logical result. However, the negative effect of increased attrition was more than offset by the increase in productivity that results from the capability to operate into short and unprepared surfaces. There are many short, unprepared surfaces in the SWA theater.

Improving Field Performance (F) when Ground Flotation (G) is at its low value results in a 50% reduction in ROT. In this case, average attrition increased from 2 to 12.5, or from 5% to 31%. When limited to long, hard runway (F-,G-) operations there was an average of 123 flights near the threat, while there was an average of 814 flights in the

vicinity of the threat when operations on short, hard runways (F+,G-) was possible. However, the greater attrition that resulted from the seven fold increase in threat exposure was not offset by an increased capability. To understand why, consider that short runways are typically soft or unprepared surfaces that can not support airlifters with poor ground flotation. If a runway is short <u>and</u> hard it is probably due to an attack. Therefore, the ability to operate on short, hard surfaces results in a greatly increased exposure to the threat. However, away from the threat there is little or no increase in productivity.

The combination of effects caused by the FG and FP interactions explains the negative coefficient for Field Performance (F). More importantly, these interactions point to the fact that survivability, field performance, and ground flotation must be considered in the context of a concept of operations. The operations concept employed in this experiment continues to assign tactical airlifters into high threat areas regardless of attrition rates. Under these conditions the advantage of being able to operate into short, unprepared fields is offset by the high attrition rates associated with many of these locations. In an actual war, measures would be taken to improve survivability if attrition rates were unacceptably high. Fighter escorts could be provided, safe corridors established, and ground times reduced. Operations into the high threat areas could be reduced or stopped altogether; one way to reduce operations would be to only transport the most critical cargo into hostile areas. Taking these measures would reduce the deleterious effects of the interactions between survivability, field performance, and ground flotation.

<u>Threat-Related Operational Control Parameters Within GAMM.</u> GAMM allows for operational control near the threat in two ways. First, there is the capability to allow the movement of only the highest

priority cargo into high threat areas. Restricting cargo movement based on priority is a function of the inflight probability of survival (P(F)) and a threshold that the user sets for each priority of cargo. If the P(F) is below the threshold, the cargo will not be shipped. The capability to restrict cargo movement based on priority was never invoked for this study. The default threshold value of .90 was used for all priorities of cargo. The lowest inflight probability of survival was .969, therefore no cargo was restricted due to a low P(F). Second, GAMM allows the user to set, as a parameter in the Runtime Command File, an attrition rate above which all operations into high threat areas will stop. Operations were not stopped near the threat for this research. Stopping operations near the threat is one way of dealing with attrition, but it may not be the most realistic approach. In actuality, cargo that could determine the outcome of the war would almost certainly be delivered. As indicated in Appendix B, invoking the option to restrict all operations into high-threat areas is not recommended.

The high attrition rates suffered in this experiment indicate that using the default and recommended values for threat related operational control parameters available with GAMM is not appropriate. However, any other values used for these parameters would be entirely arbitrary. Currently, there is no information available that relates attrition rates and threat levels to the threat-related operational parameters within GAMM. Further research needs to be done to determine appropriate values for these parameters when there is the potential for high attrition rates.

The analysis of the models developed in this study is provided in the following sections. The concept of operations used must be considered when interpreting these models. Operations into high threat areas continued regardless of the attrition rate. This policy resulted

in high attrition that offset the positive effects of increased Field Performance (F) and Ground Flotation (G).

The Analysis of the Regression Models

The following four sections discuss the regression analysis performed on each measure of effectiveness (MOE) used in this research. Each section includes a table that presents several measures of adequacy of the linear models developed for that MOE. The terms are listed in the order in which they were entered in the stepwise regression procedure. Referring to Table 6, Fleet Size (N) was the first factor entered, followed by Aircraft Size (S), and so on. Each model contains the term in that row, all terms above it, and a constant term that is tabled with the coefficients in Appendix G. Again, referring to Table 6, the model labeled as LNROT5 includes P, I, -NN, S, N, and a constant term. Quadratic terms are denoted by showing the same factor twice. For example, a quadratic term involving Fleet Size (N) is denoted by NN. Interaction terms include the designator for each factor in the interaction. For instance, the interaction between Field Performance (F) and Survivability (P) is denoted by FP.

The following three types of models are discussed for each MOE:

1. Parsimonious Model - The parsimonious model contains the fewest terms and only those terms that most impact the response.
2. Expanded Parsimonious Model - The expanded parsimonious model contains all terms in the parsimonious model.
   It also includes all significant quadratic and interaction terms that do not increase the number of characteristics presented in the parsimonious model. The expanded parsimonious provides an improvement over the parsimonious model without increasing the number of characteristics under consideration.

3. <u>Full Model</u> - The full model provides the highest precision and contains the ten most significant factors.

In the tables presented for each stepwise regression, the parsimonious model is denoted by PAR, the expanded parsimonious model by EXPAR, and the full model by FULL. The expanded parsimonious model is presented immediately after the parsimonious model so that the two models can more easily be compared.

<u>Ratio on Time (ROT)</u>

As Table 6 indicates, the parsimonious model for ROT contains just five terms, including the constant, and describes most of the variation in the responses about the mean (R^2 = .79). The mean square error (MSE) has also achieved over 70% of the reduction achieved in the full model by this point in the stepwise regression. The expanded parsimonious model consists of those terms in the parsimonious model plus the quadratic term for Aircraft Size (SS). However, as Table 6 indicates, for ROT the expanded parsimonious model does not provide much improvement over the parsimonious model. The adequacy measures for the two models are almost equal. Note that the expanded parsimonious model for ROT only contains two airlifter characteristics plus Fleet Size (N) and provides a large percentage of the predictive power contained in the full model. Clearly, of the airlifter characteristics considered, Aircraft Size (S) and Inflight Performance (I) most impact the Ratio on Time (ROT) measure of effectiveness.

There is not a large improvement between the expanded parsimonious and the full. R^2 only increases by .12 between EXPAR and FULL. The MSE does decrease by a over factor of two, but there isn't much of a penalty in terms of the standard deviation of the residuals transformed to the original data (ε_i). The standard deviation of the ε_i s only improves by .013. Assuming normality, LNROT5 will be within .08 (2 x .04) of the

Table 6. Stepwise Regression Of LNROT

Eqn ID	Added Term	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
LNROT1	N	.3976	.9643	.13413	39.60	.2066	.0654
LNROT2	S	.6485	.9792	.07960	54.42	.2066	.0544
LNROT3	-NN	.7225	.9836	.06391	50.34	.1900	.0478
PAR I	I	.7911	.9876	.04895	53.97	.1524	.0424
EXPAR	----	.8062	.9885	.04623	46.59	.1417	.0404
LNROT5	P	.8292	.9899	.04075	54.37	.1193	.0375
LNROT6	M	.8616	.9918	.03362	57.05	.1193	.0339
LNROT7	-F	.8860	.9932	.02821	59.93	.1462	.0336
LNROT8	FG	.9031	.9943	.02443	61.72	.1462	.0319
LNROT9	-SS	.9181	.9951	.02103	64.80	.1352	.0291
FULL	FP	.9283	.9958	.01878	66.02	.1034	.0274

N = Fleet Size F = Field Performance S = Aircraft Size
G = Ground Flotation I = Inflight Performance M = Maintainability
P = Survivability
$\varepsilon_i = Y_i - \hat{Y}_i$ is in terms of original data.
Max ε_i is in terms of magnitude.
Minus sign denotes a negative coefficient.

true response 95% of the time while the full model will be within .054 (2 x .027) of the response 95% of the time. Adding four terms only narrowed the 95% confidence interval by .05. However, if high precision is desired the full model would be most appropriate. Its error terms had the smallest standard deviation and it was never more than .10 from the observed response.

Ratio Delivered (RD)

As indicate in Table 7, two airlifter characteristics dominate the LNRD model. The parsimonious model contains just Aircraft Size (S), Survivability (P), and Fleet Size (N) while describing 80% of the variation in the responses about the mean. Also, 80% of the reduction

Table 7. Stepwise Regression Of LNRD

Eqn ID	Added Term	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
LNRD1	N	.3186	.9306	.06898	28.06	.2994	.1085
LNRD2	S	.5097	.9500	.05047	30.67	.2994	.0941
LNRD3	-NN	.6610	.9532	.03550	37.69	.2593	.0786
PAR P	P	.7986	.9548	.02146	56.50	.1701	.0575
EXPAR	----	.8324	.9829	.01851	45.52	.1688	.0523
LNRD5	I	.8296	.9688	.01847	54.54	-.1516	.0529
LNRD6	FG	.8542	.9707	.01610	53.71	.1270	.0501
LNRD7	-SS	.8734	.9732	.01424	53.20	.1543	.0471
LNRD8	NF	.8919	.9886	.01238	54.68	.1543	.0448
LNRD9	M	.9073	.9906	.01083	56.53	.1212	.0421
FULL	PP	.9219	.9920	.00929	60.23	.1367	.0383

N = Fleet Size F = Field Performance S = Aircraft Size
G = Ground Flotation I = Inflight Performance M = Maintainability
P = Survivability
$\varepsilon_i = Y_i - \hat{Y}_i$ is in terms of original data.
Max ε_i is in terms of magnitude.
Minus sign denotes a negative coefficient.

in MSE is achieved by this point in the stepwise regression. The expanded parsimonious model for RD contains two terms in addition to those in the parsimonious model (SS and PP) and provides slightly improved model adequacy measures. The expanded parsimonious model demonstrates that while considering just two airlifter characteristics plus Fleet Size (N) a reasonably good fit to the data can be achieved.

The return for each term added after the parsimonious model is marginal. Six terms and four additional airlifter characteristics are added between LNRD5 and the full model. This large increase in detail does not greatly improve the predictive power of the model. However, if a more accurate estimate is desired one of the last two models would be

most appropriate. Either of these models will be within .08 of the actual response 95% of the time.

Ratio of Critical Cargo Delivered (RCC)

Referring to Table 8, Fleet Size (N) plus three airlifter characteristics provide most of the descriptive power for the SQRTRCC model. The parsimonious model includes Fleet Size (N), Aircraft Size (S), Survivability (P), and Field Performance (F) while describing 78% of the variation in the responses about the mean. Also, 80% of the total reduction in MSE is achieved by this point in the stepwise regression. The expanded parsimonious model attains an R^2 of .8341 and a MSE of .00264 by including the NN, PP, FF, NF. The expanded parsimonious model demonstrates that a reasonably adequate model can be developed for RCC while considering just three airlifter characteristics plus Fleet Size (N).

While SQRTRCC9 includes four more terms than the parsimonious model, the only additional aircraft characteristic is Inflight Performance (I). Thus SQRTRCC9 presents an adequate model for RCC that involves only four airlifter characteristics plus Fleet Size (N).

Ratio of Critical Cargo Delivered (RCC) is the only MOE, other than Ratio on Time (ROT), that includes Field Performance (F) as a significant main effect. In the ROT model, F had a negative coefficient while it has a positive coefficient in the RCC model. ROT takes into account the timeliness of all cargo deliveries. As more and more aircraft are lost a larger percentage of all priorities of cargo are delivered late. However, critical cargo is delivered first and is often destined for the short fields near the FLOT. Critical cargo is least affected by aircraft attrition and most benefits from improved field performance. Attrition levels for this analysis were not high enough to cause a significant interaction between Field Performance (F) and Survivability (P) for RCC.

Table 8. Stepwise Regression Of SQRTRCC

Eqn ID	Added Term	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
SQRTRCC1	S	.2628	.9724	.01058	21.39	.3385	.1254
SQRTRCC2	N	.5041	.9814	.00723	29.99	.3385	.1066
SQRTRCC3	P	.6389	.9865	.00536	34.21	.2506	.0908
SQRTRCC4	F	.7290	.9898	.00409	38.33	.1715	.0748
PAR	-NN	.7786	.9917	.00340	39.38	.1487	.0664
EXPAR	----	.8341	.9938	.00264	38.77	.1364	.0569
SQRTRCC6	I	.8229	.9934	.00277	42.58	.1487	.0597
SQRTRCC7	PP	.8558	.9946	.00230	45.78	.1695	.0539
SQRTRCC8	FF	.8784	.9954	.00197	47.84	.1364	.0489
SQRTRCC9	NF	.8947	.9961	.00174	49.09	.1308	.0467
FULL	M	.9097	.9966	.00152	51.36	-.1142	.0436

N = Fleet Size F = Field Performance S = Aircraft Size
G = Ground Flotation I = Inflight Performance M = Maintainability
P = Survivability
$\varepsilon_i = Y_i - \hat{Y}_i$ is in terms of original data.
Max ε_i is in terms of magnitude.
Minus sign denotes a negative coefficient.

Ton Miles Per Day (Millions) (TMPD)

As Table 9 indicates, Fleet Size (N) plus two airlifter characteristics provide most of the descriptive power for TMPD. The parsimonious model includes Fleet Size (N), Aircraft Size (S), and Survivability (P) and provides an R^2 of .77. In addition, most of the reduction in MSE is achieved by this point in the stepwise regression. Finally, the parsimonious model would provide an estimate within .08 of the actual response 95% of the time. If SS and PP are added to the parsimonious model, to achieve the expanded parsimonious model, R^2

Table 9. Stepwise Regression Of SQRTTMPD

Eqn ID	Added Term	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
SQRTTMPD1	N	.3119	.9863	.00386	27.19	.1627	.0648
SQRTTMPD2	S	.5204	.9905	.00273	32.00	.1627	.0545
SQRTTMPD3	P	.6503	.9931	.00203	35.95	.1166	.0457
PAR	-NN	.7696	.9954	.00136	47.60	.0972	.0367
EXPAR	----	.8028	.9961	.00120	37.32	.0960	.0337
SQRTTMPD5	I	.8076	.9962	.00115	47.01	.0972	.0334
SQRTTMPD6	FG	.8330	.9967	.00102	45.71	.0791	.0312
SQRTTMPD7	-SS	.8537	.9971	.00091	45.00	.0734	.0291
SQRTTMPD8	NF	.8699	.9974	.00083	44.29	.0685	.0280
SQRTTMPD9	M	.8841	.9977	.00075	44.06	-.0680	.0264
FULL	PP	.8966	.9979	.00068	44.21	-.0611	.0246

N = Fleet Size F = Field Performance S = Aircraft Size
G = Ground Flotation I = Inflight Performance M = Maintainability
P = Survivability
$\varepsilon_i = Y_i - \hat{Y}_i$ is in terms of original data.
Max ε_i is in terms of magnitude.
Minus sign denotes a negative coefficient.

increases to .8076 from .7696 and MSE decreases to .00120 from .00136. The expanded parsimonious model provides aa adequate fit to the data while considering just two airlifter characteristics.

Between the parsimonious and full models four airlifter characteristics and six terms are added. These additional terms complicate the model without appreciably improving model fidelity. However, the last two models provide a 95% confidence interval of ±.05 if a high level of accuracy is desired.

Recall that the values used for TMPD in this experiment were estimates. TMPD was included in this analysis to determine if it was a suitable tactical airlift MOE for a relatively large theater of operations. This topic will be addressed in the next section.

A Common Set of Significant Characteristics

A stepwise regression provides the "best" model for a given number of terms. Another approach, which was used to develop the expanded parsimonious models in the previous sections, might be to develop the best model for a certain subset of characteristics. This approach could be used to determine the smallest number of characteristics that adequately predict the response across a group of measures.

An examination of Tables 6 through 9 reveals that Fleet Size (N), Aircraft Size (S), Inflight Performance (I), and Survivability (P) most impact airlift system performance. The first five terms entered for every MOE were, in varying orders, N, S, I, P, and NN. The one exception is RCC, where Field Performance (F) is entered two steps prior to Inflight Performance (I). Expanded parsimonious models were developed for each MOE from N, S, I, and P. The adequacy measures for these models are shown in Table 10. Each model in Table 10 includes a constant plus the following terms; N, S, I, P, NN, SS, and PP. All the models in Table 10 provide reasonably good fits to the data with the possible exception of the model for the Ratio of Critical Cargo Delivered (RCC). However, if all significant terms involving Field Performance (F) are added to the model for SQRTRCC the result is SQRTRCC9 in Table 8, a model that is more in line with the others in Table 10. Critical cargo, such as emergency resupplies and crucial unit moves, is that cargo that will have the most direct impact on the outcome of the war. Therefore, RCC would be considered a very important measure of tactical airlift system effectiveness. Due to the relative importance of RCC it would be more appropriate to include Field Performance (F) in any model used to estimate this MOE.

The models in Table 10 demonstrate that three of the four responses can be adequately characterized by considering the impact of just three airlifter characteristics plus the size of the fleet. They

Table 10. Common Expanded Parsimonious Models

Eqn ID	R^2	R^2_{unc}	MSE	F	Max ε_i	STD DEV ε_i
LNRD	.8635	.9861	.01535	48.78	.1688	.0471
LNROT	.8442	.9908	.03783	49.69	.1069	.0351
SQRTRCC	.7838	.9919	.00344	27.97	.2678	.0725
SQRTTMPD	.8408	.9968	.03151	40.75	.0960	.0300

N = Fleet Size F = Field Performance S = Aircraft Size
G = Ground Flotation I = Inflight Performance M = Maintainability
P = Survivability
Each model contains the following terms: N, S, I, P, NN, SS, and PP
$\varepsilon_i = Y_i - \hat{Y}_i$ is in terms of original data.
Max ε_i is in terms of magnitude.
Minus sign denotes a negative coefficient.

also demonstrate that the strategic MOE Ton Miles Per Day is impacted by the same airlifter characteristics as traditional tactical airlift MOEs when applied to a tactical theater that covers a large geographical region. In other words, the same set of characteristics that characterize RD, ROT, and RCC also characterize TMPD. When using GAMM in large tactical airlift theaters Ton Miles Per Day is another measure of effectiveness that can be used to determine the impact of airlifter characteristics on system performance.

Coefficients for the models in Table 10 are in Appendix G.

Maintainability and Reliability

As mentioned earlier, Reliability (R), as measured by Mean Time Between Failures (MTBFc), was dropped after the first experiment due to its insignificance. Maintainability, as measured by Mean Time To Repair (MTTR), did not have a significant impact on any of the measures of effectiveness in the Box-Behnken experiment. These findings agree with an earlier study that concluded that productivity benefits from improvements in MTTR and MTBFc above the levels currently exhibited in a

generic C-130 are relatively small (24:548). However, GAMM currently does not attrit aircraft on the ground. If aircraft were exposed to the threat while on the ground, maintainability and reliability could become more of an issue. Aircraft with a low MTBFc would be exposed to the threat more frequently. Aircraft with a high MTTR would be exposed to the threat for a longer period of time when they were out of commission. If ground survivability were modeled one might expect interactions between Survivability (P), Maintainability (M), and Reliability (R).

Comparison to Pappas's Results

Pappas developed adequate models that included only two airlifter characteristics: cabin size and ground flotation. However, his experiment was conducted in a no threat environment. The interactions between Survivability (P), Field Performance (F), and Ground Flotation (G) make it difficult to compare the results of this experiment to those of Pappas. If both experiments considered survivability and used the same concept of operations, a more useful comparison could be made. However, there are areas of the two studies that can be legitimately compared. Characteristics that were significant in this experiment, but did not significantly interact with Survivability (P), Field Performance (F), or Ground Flotation (G) can be directly compared with the same characteristics in Pappas's experiment.

Fleet Size (N) was not significant in Pappas's experiment and it is significant in all of the models developed in this analysis. Previous theater airlift studies have always shown fleet size to be very important (17:1143). Recall the range of values for Fleet Size (N) used in this experiment included an area where productivity was proportional to the number of aircraft. However, further review of the experimental setup used in Pappas's experiment indicated that the range of values used for Fleet Size (N) did not reach low enough (17:1143).

Inflight Performance (I) proved significant for all the measures

of effectiveness (MOEs) considered in this analysis. In Pappas's experiment I was not significant. The insignificance of inflight performance in Pappas's experiment is not surprising when considering that the cargo moves only 100-200 KMs in the Central American scenario. On the other hand, in Southwest Asia (SWA) 70% of the cargo is moved over 300 KMs and 35% is moved over 1000 KMs. The advantage of having a faster aircraft in SWA is obvious.

Aircraft Size (S) (Pappas's "Cabin Size") was significant for all models developed in both experiments. It was either the first or second term entered in the stepwise regression for every model developed in this experiment. The aircraft's size has a significant impact on airlift system performance across a range of scenarios.

Summary

Single degree of freedom tests for curvature on the equations developed from the 2_V^{8-2} design indicated significant curvature due to quadratic terms for each measure of effectiveness (MOE). Also, this experiment indicated that Reliability (R), when varied around that of a generic C-130, did not significantly impact airlift system performance and could be dropped from further consideration.

Due to the existence of significant quadratic terms, a Box-Behnken design was used to estimate a second-order model. A stepwise analysis of the results of the Box-Behnken experiment confirmed the existence of significant quadratic terms. Also, an unexpected result was observed. Field Performance (F) had a negative coefficient in the equations developed for Ratio on Time (ROT). A closer look, using interaction plots, revealed that very high attrition more than offset any added benefit from improved field performance. This anomaly was the result of a concept of operations that allowed flights into high threat, short field environments regardless of attrition levels.

The stepwise regression performed for each MOE was analyzed and models were developed that contained only those airlifter characteristics that were significant for all MOEs. This analysis revealed that the response surface for each MOE could be characterized reasonably well while considering only the terms including Aircraft Size (S), Survivability (P), Inflight Performance (I), and Fleet Size (N). The model developed for Ratio of Critical Cargo Delivered (RCC) was the one exception, but it could be largely improved by adding those terms that included Field Performance (F).

Maintainability (M), when changed around that of a C-130, did not have a significant effect on any of the MOEs. However, the effects of reliability and maintainability might have been different if ground survivability would have been modeled.

Finally, the results of this experiment were compared with those of Pappas. Inflight Performance (I) appears to only be significant when flight distance are relatively long. However, Aircraft (Cabin) Size (S) appears to be significant across a range of operational conditions.

VI. Conclusions and Recommendations

Summary of Analysis

The purpose of this study was to identify those tactical airlifter characteristics that most impact theater airlift system performance. Southwest Asia (SWA), a relatively large theater, was used so that comparisons could be made with an earlier study that examined the relatively small Central American (CA) theater. Using simulation and the techniques of Response Surface Methodology (RSM), metamodels were developed for each of four measures of effectiveness (MOEs). These models included only those airlifter characteristics that significantly impacted airlift system performance for the selected scenario.

A comparison of the results of this experiment with those from the CA experiment identified which airlifter characteristics were always, sometimes, and never significant. Aircraft Size (S) was significant for all MOEs in both the CA and SWA studies. Survivability (P), Field Performance (F), Ground Flotation (G), and Inflight Performance (I) were significant in some, but not all, of the models developed in the two studies. Finally, although only considered in the SWA study, Maintainability (M) and Reliability (R) never significantly impacted system performance.

In addition to identifying significant airlifter characteristics this research acted to further validate the Generalized Air Mobility Model (GAMM) by revealing some areas of the model that require improvement. Specifically, this analysis revealed that improvements need to be made in the way GAMM models the interface between survivability and the concept of operations. This study also demonstrated that the current inability to model airlifter ground survivability severely limits GAMM's usefulness for analyzing airlifter characteristics that determine the frequency and length of ground times (i.e. MTTR, MTBFc, cargo handling capabilities, etc.).

The sections that follow amplify the conclusions related to significant airlifter characteristics and GAMM validation introduced above. The chapter concludes with recommendations for further research.

Significant Airlifter Characteristics. In a tactical airlift scenario similar to Southwest Asia (SWA), characterized by long flight distances and medium to high threat levels, airlift system performance is determined primarily by Fleet Size (N) and four airlifter characteristics. The results of this study indicate that when varying airlifter characteristics around a baseline C-130, airlift system performance is primarily a function of Aircraft Size (S), Inflight Performance (I), Survivability (P), and Field Performance (F). Pappas found that in the much different Central American (CA) scenario, characterized by short flight distances and low threat levels, airlift system performance was primarily dependent on just two airlifter characteristics; Cabin Size (closely related to S) and Ground Flotation (G). Although Pappas held Survivability (P) constant at its high level for his experiment, it can be inferred that P would not be as significant in CA, where the threat is primarily due to lightly armed insurgents (7:D-2). Combining Pappas's results with those of this research reveals that across a range of operating conditions airlift system performance is most impacted by a total of five airlifter characteristics (i.e. S,I,P,F,G).

If the USAF wants the next generation tactical airlifter to perform well across a wide range of operating conditions, then design efforts would have to focus on improvements, over a baseline C-130H, in all five areas (S,I,P,F,G). However, a large fleet of such "super airlifters" might be prohibitively expensive, especially when one considers the shrinking DOD budget. It might be more practical to acquire a reduced number of the super airlifters and augment them with an aircraft designed specifically for use in low-intensity conflicts

that are confined to a relatively small geographic area, such as a Central American war. Based on Pappas's findings, something as simple as a stretch version of the C-130H with improved landing gear might greatly improve airlifter system performance in a confined, low-intensity conflict. This aircraft would be relatively inexpensive to acquire since it could be bought "off the shelf". It would also be cheap to operate since it could use much of the same maintenance and logistic support already in place for the C-130H. Finally, the modified C-130H could be used to augment the super airlifters in higher intensity conflicts by moving cargo in low threat areas. In combination, the two airlifters could possibly significantly improve airlift system performance across a wide range of operating conditions, at reduced cost.

Characteristics That Are Always Significant. As indicated in Tables 6 through 9 of Chapter V, the first two factors entered in the stepwise regression procedure were Aircraft Size (S) and Fleet Size (N) for every measure of effectiveness. This result indicates that the most important airlifter characteristic, relative to a C-130H, is Aircraft Size (S). Also, for acquisition planning purposes, the number of aircraft used largely determines how well the tactical airlift system meets demand.

The size of the aircraft was also significant in the models developed by Pappas for the Central American scenario. It can be conjectured Fleet Size (N) would also have been significant in Pappas's experiment had a wider range of values been used (17:1144).

Characteristics That Are Sometimes Significant. Table 11 summarizes the airlifter characteristics included in the parsimonious models developed for each of the measures of effectiveness used in this experiment. Table 12 contains the characteristics included in the

models developed by Pappas. Pappas did not consider the RCC and TMPD measures of effectiveness.

Table 11. Characteristics That Most Impact System Performance In SWA

Airlifter Characteristics	ROT	RD	RCC	TMPD
Aircraft Size (S)	X	X	X	X
Survivability (P)		X	X	X
Inflight Performance (I)	X			
Field Performance (F)			X	

Table 12. Characteristics That Most Impact System Performance In CA

Airlifter Characteristics	ROT	RD
Aircraft Size (S)	X	X
Ground Flotation (G)	X	X

As Tables 11 and 12 indicate, Aircraft Size (S) is the only airlifter characteristic that is always significant. To determine when other characteristics are significant both the scenario and the measure of effectiveness must be considered. For example, within the SWA scenario an aircraft designed to most improve ROT over that provided by a C-130H would include improvements in Aircraft Size (S) and Inflight Performance (I). However, by not including improvements in Survivability (P) and Field Performance (F), two airlifter characteristics that significantly impact RCC, there would not be a correspondingly significant improvement in the ratio of critical cargo delivered.

Since critical cargo is that cargo that will presumably have the most direct impact on the outcome of the war, characteristics that most impact RCC should be considered when designing the next generation tactical airlifter (22). As Table 11 indicates, improving only the three characteristics significant to RCC (S, P, and F) includes those that affect RD and TMPD when considering the SWA scenario. However, Inflight Performance (I), one of the characteristics that significantly impacts ROT, would be ignored. If the ratio of on time deliveries is considered an important measure of system performance, then improvements to I should also be considered.

This discussion on airlifter characteristics that are sometimes significant illustrates the fact that tactical airlift MOEs measure system performance at two different levels. Ratio on Time (ROT), Ratio Delivered (RD), and Millions of Ton Miles Per Day (TMPD) assess the overall performance of the tactical airlift system while Ratio of Critical Cargo Delivered (RCC) only measures the tactical airlift system's ability to deliver the highest priority cargo. The tactical airlift commander would be concerned with both levels of tactical airlift system performance. <u>Therefore, when assessing tactical airlift system performance at least two measures of effectiveness should be considered; RCC and either RCC, ROT, or TMPD.</u>

<u>Characteristics That Are Never Significant.</u> As Tables 11 and 12 indicate, Maintainability (M) and Reliability (R) are never significant. In other words, MTTR and MTBFc, when varied about the levels currently exhibited for a generic C-130, do not significantly impact system-level performance. This result is consistent with an earlier study conducted by the Aeronautical Systems Center (24:548) and implies that, within the design region described by this experiment, MTTR and MTBFc levels can be set arbitrarily, according to cost, or according to other resource limitations. Also, this result shows that larger gains in productivity

can be realized by improvements, over the baseline C-130H, in the other airlifter characteristics (S, P, F, and G) considered in this study.

Restrictions on Results. There are no standard operations concepts that relate attrition rates and threat levels to scheduling airlifters into high threat areas. This experiment tested the conjecture that ignoring attrition would not significantly affect system-level performance. Flights were scheduled into high threat areas independent of the loss rate. Unexpected interactions between Survivability (P), Field Performance (F), and Ground Flotation (G) indicate that unrestricted flights into high threat areas can result in high attrition rates that more then offset the benefits of being able to operate into short (F+), unprepared (G+) fields. These interaction terms result from the fact that there are many short, unprepared landing surfaces located near the threat in SWA. In fact, the deleterious effect of these interactions was significant enough to cause F to have a negative coefficient in the model developed for Ratio on Time (ROT).

A more realistic concept of operations would limit flights into high threat areas when attrition rates are unacceptably high. This more realistic operations concept would reduce the significance of the interactions between F, G, and P, but increase the significance of P.

The concept of operations used must be considered when interpreting the characteristics found significant in this study. For example, as Table 11 indicates, Ground Flotation (G) is not significant for any of the models developed for the SWA scenario. However, a more realistic operations concept could increase the impact of G on system-level performance.

GAMM Validation.

This research constituted a further validation effort for certain aspects of GAMM and identified some areas of the model that require

improvement. In regard to survivability and the concept of operations, there are two areas where GAMM can be improved. First, there is no threat suppression or threat evasion capability currently within GAMM. In an actual war, high priority airlift operations, such as emergency resupply or crucial unit movements, would be conducted in high threat areas even if attrition rates were high. However, threat suppression or evasion measures, such as fighter escort and the establishment of safe corridors, would be taken to help insure the success of these critical airlift operations. Currently, the inflight probability of survival (P(F)) is strictly a function of distance from the Forward Line of Troops (FLOT) and cannot be influenced by the urgency of a particular sortie. <u>Including the ability to improve the probability of survival for critical airlift operations would make GAMM a more useful tool for survivability and system-level effectiveness studies.</u>

A second area for improvement involves the way flights into high threat areas are restricted based on cargo priority. GAMM currently uses a probability of survival threshold for each cargo priority to restrict flights near the threat. If the P(F) is below the cargo's threshold, the flight is not scheduled. The ability to decide whether or not cargo will be moved based on priority <u>and</u> attrition rate is more realistic from an operational perspective. For instance, if the attrition rate was low, a commander may allow the movement of lower priority cargo near the threat. However, as more aircraft are lost, the commander will likely be more conservative with his airlift assets. <u>Operational control would be more realistically modeled if the decision to block cargo within GAMM was based on both attrition rate and cargo priority.</u>

Finally, the fact that Reliability (R) and Maintainability (M) did not significantly impact airlift system performance supports the need for continued efforts to effectively model ground survivability within

GAMM. As was pointed out in Chapter V, an aircraft with a low Mean Time Between Critical Failures (MTBFc) and a high Mean Time To Repair (MTTR) will be exposed to ground threats more often and for longer periods of time. The six hour pipeline delay that GAMM uses when aircraft fail away from their home base adds even more to ground threat exposure. If R and M had been significant in this study, it could have been inferred that they would be even more significant if ground survivability were being modeled. Since they were not significant in this study, no conclusions can be made regarding their impact had ground survivability been modeled. Studies of other airlifter characteristics related to ground time, such as cargo handling capabilities and ground servicing times, would be subject to the same limitation. GAMM will be a much more valuable analytical tool once the data required to model ground survivability is added.

Recommendations for Further Research.

A number of areas are recommended for further research.

1. More research needs to be done in regard to the relationship between Field Performance (F), Ground Flotation (G), Survivability (P), and operations concept. One approach might be to design a 2^4 factorial experiment. The airlifter characteristics F, G, and P could be varied as they were in this analysis to account for three of the four factors in the experiment. The fourth factor, the operations concept, could be changed by varying the threat-related operational control parameters within GAMM. Appropriate settings for these parameters could be coordinated through Air Mobility Command (AMC). The results of such an analysis could be used as a basis for determining just how much survivability to design into the next generation airlifter.

2. Further research could be conducted, using GAMM, to determine what mix of modified C-130Hs and super airlifters is required to meet a certain level of airlift system performance. The number of super airlifters required to meet the same level of performance could also be determined. This information could be combined with cost estimates for the two aircraft types to determine the cost savings that would result from using a mix of aircraft, if any.

3. This study, along with Pappas's thesis, and earlier studies conducted by the USAF Aeronautical Systems Center (ASC), demonstrate that using GAMM and Response Surface Methodology (RSM) leads to reasonably accurate metamodels that relate airlifter characteristics to airlift system performance. One major limitation of the metamodels developed to date is that they do not consider cost. They provide direction in terms of what characteristics to concentrate on to most improve performance, but they totally ignore the costs associated with these characteristics. While a big, fast, highly survivable airlifter capable of operating on short, unprepared surfaces may provide the most improvement in system performance, such an aircraft may be prohibitively expensive to build. Further research could use either the metamodels developed in this research or those metamodels developed with a more realistic concept of operations to assess the cost-effectiveness of the alternative configurations of interest. The metamodels could be used to measure system performance for each configuration being considered. Maximum performance for different fixed levels of cost could then be determined. Since there is more than one measure of effectiveness some sort of goal programming or weighting scheme might also be incorporated into the analysis.

4. The C-17 was not included in this study. An analysis similar to that conducted in this research could be performed to determine what characteristics are significant when the C-17 is included in the airlifter force. It might be determined that a different set of characteristics are significant when C-17 operations are included. Such a study would better define the tactical airlifter design that best complements the C-17 when the C-17 is included as part of the tactical airlift system.

Appendix A: Southwest Asia Airlift Jobs Definitions

Job No.	Description	Pr'ty 1=Hi	Freq	Wt
1	UNIT MOVE, F-15 SQDN	3	4	6T
2	UNIT MOVE, A-X SQDN	3	1	6T
3	ROUTINE RESUPPLY, POL/AMM	6	4	6T
4	UNIT MOVE, MLRS BATTERY	1	8	27T
5	PERSONNEL MOVE, DIVERTED	7	4	
6	UNIT MOVE, HAWK BATTERY	1	2	17T
7	UNIT MOVE ATK HEL BN (AH-64)	3	3	19T
8	UNIT MOVE, NBC DECON COMPANY	1	1	19T
9	ROUTINE RESUPPLY, POL/AMM	6	16	
10	AIRDROP BATTALION TASK FORCE	1	1	19T
11	AIRLAND BRIGADE	1	1	19T
12	UNIT MOVE, TACTICAL AIRLIFT SQDN	4	1	5T
13	UNIT MOVE, COMBAT ENGINEERS	4	2	20T
14	UNIT MOVE, LIGHT INFANTRY BRIGADE	1	6	7T
15	UNIT MOVEMENT, IRANIAN BRIGADE	5	2	5T
16	UNIT MOVE, MLRS BATTALION	3	1	22T
17	PERSONNEL MOVE, REPLACEMENTS	7	10	
18	EMER RESUP,AMMO/POL/FOOD/WATER	2	8	
19	EMER RESUP, PGM/POL	2	8	
20	ROUTINE RESUPPLY, PAX/REP	7	40	
21	ROUTINE RESUPPLY, RATIONS	9	12	
22	ROUTINE RESUPPLY, WATER	4	2	19T
23	UNIT MOVE, MASH	4	1	22T
24	WEAPONS DROP TO GUERRILLAS	5	2	
25	UNIT MOVE, AIR AMBULANCE CO	4	1	19T
26	MEDICAL EVACUATION	8	40	
27	EMERG RESUPPLY, ARTILLERY AMMO	2	5	
28	RETROGRADE: PARTS/EQUIP	8	6	6T
29	BACKLIFT KIA'S	9	21	
30	PERSONNEL MOVE, REPLACEMENTS	7	4	
31	UNIT MOVE, A-X WING	3	2	12T

Table adapted from one prepared by General Research Corp (7:4-7)

Appendix B: Runtime Command File Parameter Settings

Table 13. Recommended Runtime Command File Parameter Settings

DATA ITEM	USE	ALTERNATIVE VALUES	RECOMMENDED VALUE	RATIONALE
LENGTH OF SIMULATION	DEFINES THE NUMBER OF DAYS (SIMULATED TIME) BEFORE RESETTING STATISTICS	30 (DAYS)	30 (DAYS)	-LONG ENOUGH TO REMOVE TRANSIENT EFFECTS -TRACKS WITH PREVIOUS WORK
NUMBER OF REPLICATIONS	GENERATE BETTER ESTIMATES OF MEANS FOR RANDOM PROCESSES	1, 3, 5	5 (OR 1 FOR DETAILED REPORTS)	-PAST TESTING SUPPORTS FIVE
CREW DAY (HOURS)	CHECKED AT LANDING TO DETERMINE IF NEXT FLIGHT SHOULD BE HOME	10., 12., 14.	12.	-TRACKS WITH PAST ANALYSIS -ACTUAL AVERAGE CREW DAY BETWEEN 12-14 REFLECTS REASONABLE WAIVER CONDITIONS
TIME OF AVERAGE AIRBASE TEMPERATURE	SETS THE RELATIVE POSITION OF THE SINUSOIDAL TEMPERATURE CONDITIONS	10:00, 12:00	10:00	-SETTING THE AVERAGE VALUE IN LATE MORNING WILL CAUSE THE HIGHEST TEMPERATURE OF THE DAY TO OCCUR IN MID-AFTERNOON
LOWER P(s) BOUND FOR TAKEOFF ASAP	DETERMINE "THREAT"	.95, .998, 1.0	.998	-TRACKS WITH ZONE SETTING FOR NEAR FLOT (CAUSES TAKEOFF ASAP FOR NEAR FLOT AND CROSS FLOT LOCATIONS)
COMBAT FACTOR	UNDER TAKEOFF ASAP, SERVICE, LOAD, AND UNLOAD TIMES ARE DECREASED	0.3	0.3	-BLEND OF NORMAL AND COMBAT EXPERIENCE

Table 13. Recommended Runtime Command File Parameter Settings

DATA ITEM	USE	ALTERNATIVE VALUES	RECOMMENDED VALUE	RATIONALE
COMBAT MAINTENANCE FACTOR	UNDER TAKEOFF ASAP, MISSION ESSENTIAL MAINTENANCE TIME IS DECREASED.	0.3,0.5	0.5	
AIRCRAFT SCHEDULING BY INPUT (PREFERENCE) OR HIGH P(s)	DETERMINES ORDER OF CONSIDERATION OF AIRCRAFT FOR ASSIGNMENT	I, P	I (C-130 FIRST CHOICE C-17 SECOND CHOICE IN BASELINE)	-TRACKS WITH EARLIER ANALYSIS -NO EXTREME P(S) AREAS
PRIORITY THRESHOLD FOR BYPASSING NORMAL SCHEDULE	NORMAL SCHEDULING (BASED ON BACKORDER WEIGHT) WILL BE BYPASSED FOR JOBS HAVING A PRIORITY HIGHER THAN OR EQUAL TO THE THRESHOLD VALUE	1, 2, 3, 4	2	-REASONABLE CUTOFF FOR SPECIAL ATTENTION -PRODUCES REASONABLE NUMBER OF RELOCATIONS
P(S) RESTRICTED SCHEDULING	RESTRICT FLYING UNDER EXTREME CONDITIONS (P(s) OR FLEET LOSS)	Y,N	N	-NO EXTREME VALUES ARE IN USE IN P(s) ZONES
WEIGHT AND VOLUME LOAD MODULES OPTION	CONSIDERS VOLUME AS WELL AS WEIGHT IN CONDUCTING AIRCRAFT LOADING	Y,N	Y	-MOVEMENT ITEMS ARE DIMENSIONED -"CUBING OUT" IS OF CONCERN
DISTRIBUTION PARAMETER FOR ENTERING CARGO	SETS THE STEPWISE FUNCTION FOR ALLOCATING CARGO TO ALL ASSOCIATED AIRBASES BASED ON PREFERENCE AND MOG	25(TONS)	25(TONS)	-TRACKS WITH PREVIOUS ANALYSIS -CLOSE TO TYPICAL C-130 LOAD

Table 13. Recommended Runtime Command File Parameter Settings

DATA ITEM	USE	ALTERNATIVE VALUES	RECOMMENDED VALUE	RATIONALE
MEAN TIME TO ELIMINATE BLOCKED AIRCRAFT	PREVENTS AIRCRAFT FROM BEING PERMANENTLY BLOCKED	2.0 (DAYS)	2.0 (DAYS)	-TRACKS WITH PREVIOUS ANALYSIS -ALLOWS IDENT OF PROBLEM WITHOUT EXCESSIVE ANALYTICAL PENALTY
MINIMUM RAMP SPOTS FOR TRANSIT AIRCRAFT	ALLOWS NON-HOME BASED AIRCRAFT TO USE A HOME BASE	2.2, 2.0	2.0	-ALLOWS AT LEAST ONE C-17 TO OPERATE IN TRANSIT
IMMEDIATE TAKEOFF WITH RELOCATION AUTHORIZED	ALLOWS AIRCRAFT AT A BASE WITH NO BACKLOGGED CARGO TO BASES NEEDING SUPPORT AS SOON AS THE AIRCRAFT ARE READY (INSTEAD OF FLYING HOME)	Y,N	Y	-TRACKS WITH PREVIOUS STUDIES -OPERATIONALLY MORE EFFICIENT
Table adapted from a table developed by General Research Corporation and Ball Systems Engineering (6:82-84)				

Appendix C: Forward Stepwise Regression Algorithm

This algorithm was taken from Applied Linear Models by Neter and others, 1990 (15:453-454).

1. The stepwise regression routine first fits a simple linear regression model for each of the potential x (independent) variables. For each linear regression model, the F^* statistic for testing whether or not the slope is zero is obtained:

$$F_k^* = \frac{MSR(x_k)}{MSE(x_k)}$$

$MSR(x_k)$ measures the reduction in the total variation of Y associated with the use of the variable x_k. The x variable with the largest F^* value is the candidate for first addition. If this F^* value exceeds a predetermined level, the x variable is added. Otherwise, the program terminates with no x variable considered sufficiently helpful to enter the regression model.

2. Assume x_7 is the variable entered in step 1. The stepwise regression routine now fits all regression models with two x variables, where x_7 is one of the pair. For each such regression model, the partial F test statistic:

$$F_k^* = \frac{MSR(x_k|x_7)}{MSE(x_7, x_k)}$$

is obtained. this is the statistic for testing whether or not $\beta_k = 0$ when x_7 and x_k are the variables in the model. The X variable with the largest F^* value is the candidate for addition at the second stage. If this F^* value exceeds a predetermined level, the second variable is added. Otherwise the program terminates.

3. Suppose x_3 is added at the second stage. Now the stepwise regression routine examines whether any of the other x variables already

in the model should be dropped. For this example the F statistic is obtained:

$$F_7^* = \frac{MSR(x_7 \mid x_3)}{MSE(x_3, x_7)}$$

In this case there is only one such statistic. In later stages there would be a number of these F^* statistics, one for each of the variables in the model besides the one last added. The variable for which this F^* value is the smallest is the candidate for deletion. If this F^* value falls below a predetermined limit, the variable is dropped from the model. Otherwise, it is retained.

4. Suppose x_7 is retained so that both x_3 and x_7 are now in the model. The stepwise routine now examines which x variable is the next candidate to enter, then examines whether any variables already in the model should be dropped, and so on until no further x variables can either be added or deleted, at which point the search terminates.

Appendix D: Factor Levels

Table 14. High (+) and Low (-) Factor Levels

Field Performance (F)

High Level (+)	- SEA LEVEL - HOT 103 F	COLD 59 F	- 5000 FEET- HOT 84 F	COLD 41 F
CTOL TO AT MAX USEFUL LOAD(FT)	2573	1910	2916	2457
CTOL LD AT MAX USEFUL LOAD(FT)	1809	1695	2052	1920
CTOL TO AT MID USEFUL LOAD(FT)	1374	1106	1598	1374
CTOL LD AT MID USEFUL LOAD(FT)	1494	1400	1668	1557
CTOL TO AT ZERO USEFUL LOAD(FT)	1159	1079	1272	1180
CTOL LD AT ZERO USEFUL LOAD(FT)	653	603	775	696

Low Level (-)				
CTOL TO AT MAX USEFUL LOAD(FT)	5107	3791	5788	4877
CTOL LD AT MAX USEFUL LOAD(FT)	3591	3365	4074	3810
CTOL TO AT MID USEFUL LOAD(FT)	2726	2195	3172	2727
CTOL LD AT MID USEFUL LOAD(FT)	2966	2780	3310	3091
CTOL TO AT ZERO USEFUL LOAD(FT)	2301	2141	2526	2343
CTOL LD AT ZERO USEFUL LOAD(FT)	1297	1197	1537	1382

Aircraft Size (S)	High Level(+)	Low Level(-)
CARGO BAY WIDTH (INCH)	136	104
CARGO BAY HEIGHT (INCH)	120	97
CARGO BAY LENGTH (INCH)	782	391
CARGO BAY DOOR WIDTH (INCH)	136	104
CARGO BAY DOOR HEIGHT (INCH)	120	97
CTOL MAX USEFUL LOAD (LBS)	98420	49580
CTOL MID USEFUL LOAD (LBS)	51870	26130
MAXIMUM CABIN PAYLOAD (LBS)	55000	24198
MAXIMUM FERRY FUEL (LBS)	80000	40300
CRUISE FUEL (LBS/HR)	7050	3550
AIRCRAFT SPOT FACTOR (NO)	1.3	.7
CARGO THRESHOLD FOR RELOCATION(LBS)	12000	6400

Inflight Performance (I)	High Level(+)	Low Level(-)
CRUISE SPEED (KNOTS)	314	156

Ground Flotation/Wheel Loading (G)	High Level(+)	Low Level(-)
LCN - MAX USEFUL LOAD (NO)	21	43
LCN - AT ZERO USEFUL LOAD (NO)	6	12

Fleet Size (N)	20	60

Table 14. High (+) and Low (+) Factor Levels

	High Level(+)	Low Level(-)
Reliability (R)		
MTBFc (HRS)	2	6
Maintainability (M)		
MTTR (HRS)	0.6	3

Survivability (P)
Low Level (-)

Zone	Minimum Distance to Flot (KM)	Maximum Distance to Flot (km)	Probability of Survival Factor	Conditional Probability no Battle Damage
1	10	9999	1.000	1.000
2	0	10	0.983	0.995
3	-50	0	0.969	0.915
4	-9999	-50	0.970	0.918

High Level (+)
All probabilities set to 1.

Appendix E: Results of the 2- Level Fractional Factorial Experiment

Table 15. 2 - Level Fractional Factorial Experimental Results

Run No	Factor Settings N F S G I R M P	Ratio Delivered	Ratio on Time	Ratio of Critical Cargo Delivered	Million Ton Miles Per Day
1	− − − − − − + +	.2119	.0559	.1824	.1479
2	+ − − − − − − −	.2925	.0929	.1693	.2132
3	− + − − − − − −	.1032	.0241	.0542	.0698
4	+ + − − − − + +	.3921	.1131	.3773	.2304
5	− − + − − − − +	.2995	.0679	.3089	.2099
6	+ − + − − − + −	.4819	.2675	.3627	.3507
7	− + + − − − + −	.1398	.0475	.1319	.0888
8	+ + + − − − − +	.4499	.1586	.4966	.2701
9	− − − + − − − +	.1548	.0391	.1253	.1090
10	+ − − + − − + −	.3379	.1313	.1934	.2403
11	− + − + − − + −	.0951	.0247	.0685	.0655
12	+ + − + − − − +	.4279	.1295	.4334	.2561
13	− − + + − − + +	.4302	.1211	.4203	.2960
14	+ − + + − − − −	.4532	.2261	.3263	.3261
15	− + + + − − − −	.2509	.0824	.2316	.1580
16	+ + + + − − + +	.8317	.4049	.9571	.4889
17	− − − − + − + −	.2341	.0699	.1414	.1744
18	+ − − − + − − +	.4458	.1789	.3257	.3159
19	− + − − + − − +	.1661	.0485	.1811	.0977
20	+ + − − + − + −	.2463	.0818	.2625	.1476
21	− − + − + − − −	.3607	.1623	.2868	.2670
22	+ − + − + − + +	.6674	.4242	.5907	.4571
23	− + + − + − + +	.4754	.1443	.6796	.2717
24	+ + + − + − − −	.4324	.1484	.5064	.2644
25	− − − + + − − −	.2019	.0527	.1194	.1520
26	+ − − + + − + +	.4953	.2502	.3633	.3476
27	− + − + + − + +	.3693	.1141	.4229	.2249

Table 15. 2 - Level Fractional Factorial Experimental Results

Run No	Factor Settings N F S G I R M P	Ratio Delivered	Ratio on Time	Ratio of Critical Cargo Delivered	Million Ton Miles Per Day
28	+ + - + + - - -	.4326	.1332	.4406	.2707
29	- - + + + - + -	.4082	.2029	.3354	.2948
30	+ - + + + - - +	.6375	.3731	.5444	.4441
31	- + + + + - - +	.5497	.1699	.7774	.3158
32	+ + + + + - + -	.8125	.4577	.9309	.4965
33	- - - - - + + -	.2025	.0632	.1157	.1455
34	+ - - - - + - +	.4028	.1338	.2938	.2795
35	- + - - - + - +	.1342	.0407	.1375	.0822
36	+ + - - - + + -	.2854	.0729	.2762	.1623
37	- - + - - + - -	.3152	.1164	.2491	.2302
38	+ - + - - + + +	.6366	.3597	.5377	.4361
39	- + + - - + + +	.2839	.0827	.3270	.1664
40	+ + + - - + - -	.3830	.1174	.4356	.2342
41	- - - + - + - -	.1639	.0462	.0875	.1170
42	+ - - + - + + +	.4416	.1547	.3216	.3059
43	- + - + - + + +	.2556	.0728	.2990	.1475
44	+ + - + - + - -	.3641	.0883	.3814	.2141
45	- - + + - + + -	.3406	.1334	.2665	.2468
46	+ - + + - + - +	.6143	.3022	.5050	.4265
47	- + + + - + - +	.4244	.1101	.5738	.2462
48	+ + + + - + + -	.7594	.3254	.8844	.4458
49	- - - - + + + +	.3446	.1093	.2866	.2450
50	+ - - - + + - -	.3643	.1727	.2070	.2636
51	- + - - + + - -	.0875	.0277	.0627	.0606
52	+ + - - + + + +	.5114	.2417	.5054	.3212
53	- - + - + + - +	.5256	.1979	.5074	.3689
54	+ - + - + + + -	.5029	.3922	.3836	.3650
55	- + + - + + + -	.1874	.0710	.1685	.1243

Table 15. 2 - Level Fractional Factorial Experimental Results

Run No	Factor Settings N F S G I R M P	Ratio Delivered	Ratio on Time	Ratio of Critical Cargo Delivered	Million Ton Miles Per Day
56	+ + + - + + - +	.7271	.3712	.8797	.4318
57	- - - + + + - +	.2843	.0744	.2495	.2041
58	+ - - + + + + -	.3828	.2070	.2371	.2792
59	- + - + + + + -	.2058	.0492	.1873	.1352
60	+ + - + + + - +	.5823	.3060	.5428	.3675
61	- - + + + + + +	.6047	.2923	.5505	.4202
62	+ - + + + + - -	.4960	.3487	.3762	.3596
63	- + + + + + - -	.3036	.0953	.3806	.1819
64	+ + + + + + + +	.8930	.6755	.9717	.5319
65	0 0 0 0 0 0 0 0	.4429	.1828	.3793	.2892
66	0 0 0 0 0 0 0 0	.4301	.1828	.3701	.2869
67	0 0 0 0 0 0 0 0	.4287	.1806	.3527	.2818
68	0 0 0 0 0 0 0 0	.4295	.1726	.3640	.2865

This 2_V^{8-2} design was taken from Empirical Model-Building and Response Surfaces (1;163)

Appendix F: Results of the Box-Behnken Experiment

Table 16. Box-Behnken Experimental Results

Run No	Factor Settings N F S G I M P	Ratio Delivered	Ratio on Time	Ratio Critical Cargo Delivered	Million Ton Miles Per Day
1	0 0 0 − − − 0	.3519	.1133	.2577	.2369
2	0 0 0 + − − −	.3614	.1189	.2574	.2457
3	0 0 0 − + − 0	.4425	.1811	.3931	.2890
4	0 0 0 + + − 0	.4415	.1794	.3905	.2915
5	0 0 0 − − + 0	.4239	.1692	.3340	.2758
6	0 0 0 + − + 0	.4302	.1701	.3380	.2822
7	0 0 0 − + + 0	.4891	.2366	.4419	.3160
8	0 0 0 + + + 0	.4814	.2402	.4310	.3138
9	− 0 0 0 0 − −	.2190	.0737	.1535	.1530
10	+ 0 0 0 0 − −	.4531	.1974	.4030	.2987
11	− 0 0 0 0 + −	.2477	.0849	.1934	.1717
12	+ 0 0 0 0 + −	.4945	.2507	.4399	.3199
13	− 0 0 0 0 − +	.2850	.0789	.2888	.1868
14	+ 0 0 0 0 − +	.5739	.2337	.5543	.3653
15	− 0 0 0 0 + +	.3648	.1057	.3849	.2357
16	+ 0 0 0 0 + +	.6288	.3181	.5776	.4009
17	0 − 0 0 − 0 −	.3591	.1567	.2874	.2611
18	0 + 0 0 − 0 −	.3094	.0994	.3101	.1929
19	0 − 0 0 + 0 −	.3985	.2210	.3436	.2960
20	0 + 0 0 + 0 −	.4280	.1468	.4742	.2683
21	0 − 0 0 − 0 +	.5078	.1720	.4137	.3509
22	0 + 0 0 − 0 +	.5377	.1763	.5730	.3196
23	0 − 0 0 + 0 +	.5984	.3123	.4973	.4224
24	0 + 0 0 + 0 +	.6943	.3639	.6954	.4312
25	− − 0 − 0 0 0	.3110	.1137	.2600	.2277
26	+ − 0 − 0 0 0	.4057	.2406	.3550	.3006
27	− + 0 − 0 0 0	.1667	.0610	.1632	.1079
28	+ + 0 − 0 0 0	.3769	.1228	.4707	.2377
29	− − 0 + 0 0 0	.3123	.1174	.2607	.2290

Table 16. Box-Behnken Experimental Results

Run No	Factor Settings N F S G I M P	Ratio Delivered	Ratio on Time	Ratio Critical Cargo Delivered	Million Ton Miles Per Day
30	+ − 0 + 0 0 0	.4043	.2375	.3556	.2999
31	− + 0 + 0 0 0	.3116	.0967	.3443	.1973
32	+ + 0 + 0 0 0	.6017	.3052	.6590	.4021
33	0 0 − − 0 0 −	.2773	.0964	.1934	.1929
34	0 0 + − 0 0 −	.4320	.2032	.4272	.2806
35	0 0 − + 0 0 −	.2829	.1007	.2048	.1958
36	0 0 + + 0 0 −	.4324	.1968	.4188	.2869
37	0 0 − − 0 0 +	.4162	.1202	.3370	.2590
38	0 0 + − 0 0 +	.6381	.2543	.6432	..3986
39	0 0 − + 0 0 +	.4139	.1195	.3356	.2595
40	0 0 + + 0 0 +	.6357	.2560	.6339	.4005
41	− 0 − 0 − 0 0	.1621	.0468	.0930	.1128
42	+ 0 − 0 − 0 0	.2959	.1169	.2272	.2010
43	− 0 + 0 − 0 0	.2683	.0914	.2062	.1824
44	+ 0 + 0 − 0 0	.4509	.2209	.4507	.2976
45	− 0 − 0 + 0 0	.2144	.0635	.1496	.1521
46	+ 0 − 0 + 0 0	.3523	.1648	.3054	.2331
47	− 0 + 0 + 0 0	.2853	.1077	.2592	.1883
48	+ 0 + 0 + 0 0	.4881	.2870	.4843	.3393
49	0 − − 0 0 − 0	.3171	.1152	.1719	.2305
50	0 + − 0 0 − 0	.3112	.0853	.2812	.1895
51	0 − + 0 0 − 0	.4831	.2746	.3587	.3452
52	0 + + 0 0 − 0	.6043	.2359	.6493	.3632
53	0 − − 0 0 + 0	.3538	.1531	.1991	.2563
54	0 + − 0 0 + 0	.3675	.1080	.3661	.2146
55	0 − + 0 0 + 0	.4940	.3224	.3749	.3545
56	0 + + 0 0 + 0	.6690	.3003	.7290	.4047
57	0 0 0 0 0 0 0	.4323	.1777	.3712	.2856
58	0 0 0 0 0 0 0	.4313	.1784	.3594	.2867

Table 16. Box-Behnken Experimental Results

Run No	Factor Settings N F S G I M P	Ratio Delivered	Ratio on Time	Ratio Critical Cargo Delivered	Million Ton Miles Per Day
59	0 0 0 0 0 0 0	.4256	.1739	.3545	.2826
60	0 0 0 0 0 0 0	.4214	.1663	.3503	.2805
61	0 0 0 0 0 0 0	.4378	.1825	.3702	.2912
62	0 0 0 0 0 0 0	.4286	.1750	.3797	.2847

This Box-Behnken Design was taken from Empirical Model-Building and Response Surfaces. (1;519)

Appendix G: Model Coefficients

Table 17. Coefficients For LNROT Models

Eqn ID	Coefficients					
	CONST	N	S	NN	I	P
LNROT1	-1.84973	.47043				
LNROT2	-1.84973	.47043	.37370			
LNROT3	-1.74933	.47043	.37370	-.25936		
PAR4	-1.74933	.47043	.37370	-.25936	.19542	
EXPAR	-1.69989	.47043	.37370	-.26966	.19542	
LNROT5	-1.74933	.47043	.37370	-.25936	.19542	.14554
LNROT6	-1.74933	.47043	.37370	-.25936	.19542	.14554
LNROT7	-1.74933	.47043	.37370	-.25936	.19542	.14554
LNROT8	-1.74933	.47043	.37370	-.25936	.19542	.14554
LNROT9	-1.69989	.47043	.37370	-.26966	.19542	.14554
FULL	-1.69989	.47043	.37370	-.26966	.19542	.14554
Eqn ID	Coefficients					
	M	F	FG	SS	FP	
EXPAR				.11741		
LNROT6	.13424					
LNROT7	.13424	-.11656				
LNROT8	.13424	-.11656	.16901			
LNROT9	.13424	-.11656	.16901	-.11741		
FULL	.13424	-.11656	.16901	-.11741	.13023	

N = Fleet Size B = Ground Flotation S = Aircraft Size
F = Field Performance I = Inflight Performance M = Maintainability
P = Survivability

Table 18. Coefficients For LNRD Models

Eqn ID	Coefficients					
	CONST	N	S	NN	P	I
LNRD1	-.92924	.28396				
LNRD2	-.92924	.28396	.21993			
LNRD3	-.83250	.28396	.21993	-.24991		
PAR	-.83250	.28396	.21993	-.24991	.18661	
EXPAR	-.83112	.28396	.21993	-.25019	.18661	
LNRD5	-.83250	.28396	.21993	-.24991	.18661	.08866
LNRD6	-.83250	.28396	.21993	-.24991	.18661	.08866
LNRD7	-.79491	.28396	.21993	-.25774	.18661	.08866
LNRD8	-.79491	.28396	.21993	-.25774	.18661	.08866
LNRD9	-.79491	.28396	.21993	-.25774	.18661	.08866
FULL	-.83112	.28396	.21993	-.25019	.18661	.08866
Eqn ID	Coefficients					
	FG	SS	NF	M	PP	
EXPAR		-.08173			.07845	
LNRD6	.13657					
LNRD7	.13657	-.08927				
LNRD8	.13657	-.08927	.11872			
LNRD9	.13657	-.08927	.11872	.06232		
FULL	.13657	-.08173	.11872	.06232	.07845	

N = Fleet Size B = Ground Flotation S = Aircraft Size
F = Field Performance I = Inflight Performance M = Maintainability
P = Survivability

Table 19. Coefficients For SQRTRCC Models

Eqn ID	Coefficients					
	CONST	S	N	P	F	NN
SQRTRCC1	.59738	.09712				
SQRTRCC2	.59738	.09712	.09306			
SQRTRCC3	.59738	.09712	.09306	.06954		
SQRTRCC4	.59738	.09712	.09306	.06954	.05685	
PAR	.61824	.09712	.09306	.06954	.05685	-.05387
EXPAR	.58275	.09712	.09306	.06954	.05685	-.04648
SQRTRCC6	.61824	.09712	.09306	.06954	.05685	-.05387
SQRTRCC7	.59968	.09712	.09306	.06954	.05685	-.05001
SQRTRCC8	.58275	.09712	.09306	.06954	.05685	-.04648
SQRTRCC9	.58275	.09712	.09306	.06954	.05685	-.04648
FULL	.58275	.09712	.09306	.06954	.05685	-.04648
Eqn ID	Coefficients					
	I	PP	FF	NF	M	
EXPAR		.04759	.03667			
SQRTRCC6	.03987					
SQRTRCC7	.03987	.04407				
SQRTRCC8	.03987	.04759	.03667			
SQRTRCC9	.03987	.04759	.03667	.04193		
FULL	.03987	.04759	.03667	.04193	.02318	

N = Fleet Size B = Ground Flotation S = Aircraft Size
F = Field Performance I = Inflight Performance M = Maintainability
P = Survivability

Table 20. Coefficients For SQRTTMPD Models

Eqn ID	Coefficients					
	CONST	N	S	P	NN	I
SQRTTMPD1	.51807	.06616				
SQRTTMPD2	.51807	.06616	.05409			
SQRTTMPD3	.51807	.06616	.05409	.04271		
PAR	.53830	.06616	.05409	.04271	-.05226	
EXPAR	.53962	.06616	.05409	.04271	-.05254	
SQRTTMPD5	.53830	.06616	.05409	.04271	-.05226	.02309
SQRTTMPD6	.53830	.06616	.05409	.04271	-.05226	.02309
SQRTTMPD7	.54751	.06616	.05409	.04271	-.05418	.02309
SQRTTMPD8	.54751	.06616	.05409	.04271	-.05418	.02309
SQRTTMPD9	.54751	.06616	.05409	.04271	-.05418	.02309
FULL	.53962	.06616	.05409	.04271	-.05254	.02309
Eqn ID	Coefficients					
	FG	SS	NF	M	PP	
EXPAR		-.02021			.01707	
SQRTTMPD6	.03268					
SQRTTMPD7	.03268	-.02185				
SQRTTMPD8	.03268	-.02185	.02612			
SQRTTMPD9	.03268	-.02185	.02612	.01410		
FULL	.03268	-.02021	.02612	.01410	.01707	

N = Fleet Size B = Ground Flotation S = Aircraft Size
F = Field Performance I = Inflight Performance M = Maintainability
P = Survivability

Table 21. Common Expanded Parsimonious Coefficients for LNRD and LNROT

LNRD		LNROT	
CONSTANT	-.83112	CONSTANT	-1.71070
N	.28396	N	.47043
S	.21993	S	.37370
NN	-.25019	NN	-.26741
P	.18661	I	.19542
I	.08866	SS	-.11516
SS	-.08173	P	.14554
PP	.07845	PP	.02341
N = Fleet Size I = Inflight Performance		S = Aircraft Size P = Survivability	

Table 22. Common Expanded Parsimonious Coefficients for SQRTRCC and SQRTTMPD

SQRTRCC		SQRTTMPD	
CONSTANT	.61485	CONSTANT	.53962
N	.09306	N	.06616
S	.09712	S	.05409
P	-.05317	P	.04271
NN	.03987	NN	-.05254
I	-.03287	I	.02309
PP	.06954	PP	.01707
SS	.04091	SS	-.02021
N = Fleet Size I = Inflight Performance		S = Aircraft Size P = Survivability	

Appendix H: Correlation Matrices

Table 23. Correlation Matrix for LNROT

	LNROT	N	S	NN	I	P	M
N	0.6305						
S	0.5009	0.0000					
NN	-0.2722	-0.0000	0.0000				
I	0.2619	-0.0000	-0.0000	0.0000			
P	0.1951	0.0000	-0.0000	0.0000	-0.0000		
M	0.1799	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	
F	-0.1562	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000
FG	0.1308	-0.0000	0.0000	-0.0000	0.0000	-0.0000	0.0000
SS	-0.0984	-0.0000	-0.0000	-0.0877	-0.0000	0.0000	-0.0000
FP	0.1008	0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000

	F	FG	SS
FG	0.0000		
SS	0.0000	-0.0000	
FP	0.0000	0.0000	0.0000

Table 24. Correlation Matrix for LNRD

	LNRD	N	S	NN	P	I	FG
N	0.5645						
S	0.4372	0.0000					
NN	-0.3889	-0.0000	0.0000				
P	0.3709	0.0000	-0.0000	0.0000			
I	0.1762	-0.0000	-0.0000	0.0000	-0.0000		
FG	0.1567	-0.0000	0.0000	-0.0000	-0.0000	0.0000	
SS	-0.1037	-0.0000	-0.0000	-0.0877	0.0000	-0.0000	-0.0000
NF	0.1363	-0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000
M	0.1239	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000
PP	0.1674	-0.0000	0.0000	-0.0877	0.0000	0.0000	0.0000

	SS	NF	M
NF	-0.0000		
M	-0.0000	0.0000	
PP	-0.0877	0.0000	0.0000

Table 25. Correlation Matrix for SQRTRCC

	SQRTRCC	S	N	P	F	NN	I
S	0.5127						
N	0.4913	0.0000					
P	0.3671	-0.0000	0.0000				
F	0.3001	-0.0000	-0.0000	-0.0000			
NN	-0.2227	0.0000	-0.0000	0.0000	-0.0000		
I	0.2105	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	
PP	0.2003	0.0000	-0.0000	0.0000	0.0000	-0.0877	0.0000
FF	0.1511	-0.0000	-0.0000	0.0000	-0.0000	-0.0877	-0.0000
NF	0.1278	0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000
M	0.1224	-0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000

	PP	FF	NF
FF	-0.0877		
NF	0.0000	-0.0000	
M	0.0000	-0.0000	0.0000

Table 26. Correlation Matrix for SQRTTMPD

	SQRTTMPD	N	S	P	NN	I	FG
N	0.5584						
S	0.4566	0.0000					
P	0.3605	0.0000	-0.0000				
NN	-0.3454	-0.0000	0.0000	0.0000			
I	0.1949	-0.0000	-0.0000	-0.0000	0.0000		
FG	0.1593	-0.0000	0.0000	-0.0000	-0.0000	0.0000	
SS	-0.1130	-0.0000	-0.0000	0.0000	-0.0877	-0.0000	-0.0000
NF	0.1273	-0.0000	0.0000	0.0000	-0.0000	0.0000	0.0000
M	0.1191	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	0.0000
PP	0.1550	-0.0000	0.0000	0.0000	-0.0877	0.0000	0.0000

	SS	NF	M
NF	-0.0000		
M	-0.0000	0.0000	
PP	-0.0877	0.0000	0.0000

Bibliography

1. Box, George and Norman Draper. Empirical Model-Building and Response Surfaces. New York: John Wiley and Sons, Inc., 1987.

2. Box, George and others. Statistics for Experimenters. New York: John Wiley and Sons, Inc., 1978.

3. Bryant, Capt Joseph C. and Capt Stephen R. Gorden. User Need Satisfaction as a Basis For Tactical Airlift Scheduling. MS Thesis GST/OS/84M-4. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 1984 (AD-A141149).

4. Cornell, John. How to Apply Response Surface Methodology. Milwaukee: American Society for Quality Control, 1990.

5. Cornell, John and Andre Khuri. Response Surfaces: Designs and Analysis. New York: Marcel Dekker Inc. Press, 1987.

6. Future Theater Airlift Studies Southwest Asian Scenario: Generalized Air Mobility Model (GAMM) Data File Generation: Final Report Draft, Directorate of Advanced Systems Analysis, ASC/XRM, Aeronautical Systems Center. Contract F33657-90-D-2189 Task 0008 with General Research Corporation and Ball Systems Engineering Division. Wright-Patterson AFB OH, August 1992.

7. Future Theater Airlift Studies Southwest Asian Scenario Jobs Definitions: Final Report, Directorate of Mission Area Planning, ASD/XRS, Aeronautical Systems Division. Contract F33657-90-D-2189 Task 0008 with General Research Corporation and Ball Systems Engineering Division. Wright-Patterson AFB OH, March 1992.

8. Generalized Air Mobility Model (GAMM) Release 3.0 Programmer/Analyst's Manual, Directorate of Advanced Systems Analysis, ASD/XRM, Aeronautical Systems Division. Contract F33657-86-D-0157/0014 with General Research Corporation. Wright-Patterson AFB OH, October 1989.

9. Generalized Air Mobility Model (GAMM) Release 3.0 User's Manual, Directorate of Advanced Systems Analysis, ASD/XRM, Aeronautical Systems Division. Contract F33657-86-D-0157/0014 with General Research Corporation. Wright-Patterson AFB OH, October 1989.

10. Knowles, Brig Gen Billy M. "Tactical Airlift," Airpower Journal, 1:41-47 (Fall 1987).

11. Law, Averill and David Kelton. Simulation Modeling and Analysis. New York: McGraw Hill Inc., 1991.

12. Montgomery, D. C. Design and Analysis of Experiments. New York: John Riley and Sons Inc., 1991.

13. Mason, robert and others. Statistical Design and Analysis of Experiments. New York: John Wiley and Sons Inc., 1989.

14. Morse, Phillip and George Kimball. Methods of Operations Research. Los Altos: Peninsula Publishing, 1970.

15. Neter, John and others. Applied Linear Statistical Models. Boston: Richard D. Irwin Inc., 1990.

16. Pappas, Paul. An Analysis of Fixed Wing Tactical Airlifter Characteristics Using An Intra-Theater Airlift Computer Model. MS Thesis GLM/ENS/91S-50. School of Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991.

17. Pappas, Paul and others. "Impact of Theater Airlifter Characteristics On Future Theater Airlift System Productivity," Proceedings of the 1992 Winter Simulation Conference. 1139-1144. Washington D.C.: Society for Computer Simulation, 1992.

18. R & M Inputs. Directorate of Manpower, Personnel, and Training; Deputy Chief of Staff for Acquisition Logistics; Aeronautical Systems Division. Letter of Review for Reliability and Maintainability Inputs for the Generalized Air Mobility Model (GAMM). Wright-Patterson AFB OH, 1990.

19. Sargent, Robert. "A Tutorial on Validation and Verification of Simulation Models," Proceedings of the 1988 Winter Simulation Conference. 33-39. San Diego: Society for Computer Simulation, 1989

20. Shine, Col Alexander P. "Tactical Mobility - An Airlift Perspective," Proceedings of the Military Operations Research Society Mini-Symposium:Analysis of Tactical Transportation: Progress and Challenges (TACTRAN). 5:1-26. Fort Belvoir VA: Defense Systems Management College, 1988 (AD-A217344).

21. Whitner, Richard B. and Osman Balci. "Guidelines for Selecting and Using Simulation Model Verification Techniques," Proceedings of the 1989 Winter Simulation Conference. 559-568. San Diego: Society for Computer Simulation, 1989.

22. Wourms, Steven. Chief of Assessments Section Integration Branch, Studies and Analysis Division Development Planning Directorate. Personal Interview. Aeronautical Systems Center, Wright-Patterson AFB OH, 15 September 1992.

23. Wourms, Steven. Chief of Assessments Section Integration Branch, Studies and Analysis Division Development Planning Directorate. Personal Interview. Aeronautical Systems Center, Wright-Patterson AFB OH, 2 February 1992.

24. Wourms, Steven and Kenneth Bauer. "Impact of C-130 Reliability and Maintainability on Future Theater Airlift System Productivity," Proceedings of the 1989 Winter Simulation Conference. 542-548. Phoenix: Society for Computer Simulation, 1991.

Vita

Captain John Koger was born in Cincinnati, Ohio on 10 December 1960. He graduated high school in the spring of 1978 and entered the University of Cincinnati in the fall of that year. He graduated the University of Cincinnati in December of 1983 receiving a Bachelor of Science degree in Mathematics and a reserve commission in the USAF. He then entered pilot training and received his wings in March of 1985. His first flying assignment was with the 10th Military Airlift Squadron (MAS) in Zweibrucken, Germany, where he flew the C-23A, a small two engine transport. In May of 1988 he was assigned to the 7th MAS at Travis AFB in California, where he flew the C-141B. In August of 1991 Captain Koger entered the School of Engineering, Air Force Institute of Technology.

Permanent Address: 1009 Cardinal Dr.,
Enon, Ohio 45323

AIR FORCE INSTITUTE OF TECHNOLOGY

SABRE::JKOGER

JOB 219

THESIS

Owner UIC: **[STUDENT,JKOGER]**
Account: **GST93J**

Priority:	100
Submit queue:	PS_LPS20
Submitted:	27-APR-1993 15:16
Printer queue:	LPSLASER
Printer device:	LPS20
Started:	27-APR-1993 15:31

Digital Equipment Corporation **PrintServer 20 LPS20**

VAX/VMS Version A5.5-2 DECprint Printing Services V4.1

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
March 1993

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
An Analysis of Tactical Airlifter Characteristics and Their Impact on Theater Airlift System Performance

5. FUNDING NUMBERS

6. AUTHOR(S)
John Koger

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GST/ENS/93M-05

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This study used computer simulation and Response Surface Technology to determine what tactical airlifter characteristics most impact theater airlift system performance in a Sowthwest Asia (SWA) scenario. Some aircraft characteristics were grouped into funtional sets, while others were considered individually. After screening one characteristic, reliability, with a two-level factorial experiment, a Box-Behnken design was used to estimate second-order metamodels. A stepwise regression procedure indicated that, if attrition rates are ignored, airlift system performance is most impacted by the aircraft's size, survivability, cruise speed, and ability to operate on short fields. The SWA scenario used in this study covers a large geographical area and varying threat levels and types. The results of this study were compared with those of an earlier study that used the much smaller, low threat Central American scenario. It was determined that across a range of scenarios airlift system performance is most affected by the aircraft's size, survivability, cruise speed, ability to operate on short fields, and ability to operate on unprepared surfaces.

14. SUBJECT TERMS
theater airlift, tactical airlift

15. NUMBER OF PAGES
112

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19 SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)